

Diagonal-unitary 2-designs and their implementations by quantum circuits

Yoshifumi Nakata¹ and Mio Murao^{1,2}

¹ *Department of Physics, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan*

² *Institute for Nano Quantum Information Electronics, University of Tokyo, Tokyo 153-8505, Japan*

(Dated: March 8, 2013)

We study *random diagonal-unitary matrices*, an ensemble of unitary matrices diagonal in a given basis with randomly distributed phases for their eigenvalues. We investigate how to efficiently implement random diagonal-unitary matrices by quantum circuits and introduce *diagonal-unitary t -designs* that simulate diagonal-unitary matrices up to the t -th order moments. We present two efficient implementations of diagonal-unitary 2-designs in the computational basis by using quantum circuits. One is composed of single-qubit diagonal gates and controlled-phase gates with randomized phases, which achieves an exact diagonal-unitary 2-design after applying $O(N^2)$ gates. If the controlled-Z gates are used instead of controlled-phase gates, the circuit achieves an approximate 2-design and it is necessary to apply $O(N^2(N + \log 1/\epsilon))$ gates. We also provide applications of the diagonal-unitary 2-designs implemented by the quantum circuits, a generation of a 2-design of random states and a quantum simulator of canonical states of classical Hamiltonians.

PACS numbers: 03.67.Ac, 03.67.Bg, 05.67.-a

I. INTRODUCTION

Random unitary matrices are an ensemble of unitary matrices uniformly distributed in terms of the Haar measure on the unitary group. They have been studied in the context of random matrix theories [1] and are recently used for applications in quantum information processing [2–9]. Since random unitary matrices provide a method of random sampling of quantum states, they are also used for investigating typical properties of physical systems when details of the system are unknown [10, 11]. An ensemble of quantum states generated by random unitary matrices are known as random states and have been particularly investigated for studying typical properties of quantum states [12–26].

Efficient implementations of random unitary matrices are necessary to exploit them for quantum information processing [27, 28]. In particular, the implementations by quantum circuits are intensely studied [4, 8, 29–31]. It has been shown that the ensemble of unitary matrices simulating up to the t -th order of statistical moments of random unitary matrices, which is referred to as *random unitary t -design*, can be approximately achieved by quantum circuits. The number of elementary gates in the circuits scales with $O(Nt^4 \log t(Nt + \log 1/\epsilon))$, where N is the system size and ϵ is an allowed error for the approximation [31]. This result implies that an approximate t -design of random states can be efficiently generated by the quantum circuit.

Besides the applications for quantum information processing, the random unitary matrices provide a method to investigate typical dynamics in uniform systems. However, it is often the case that we would like to investigate dynamics under some constraints of systems. This leads to an idea to investigate an ensemble of unitary matrices with some constraints. For instance, when the system has a symmetry, the Hilbert space is decomposed into irreducible subspaces and it is natural to consider

dynamics in each subspace separately. Here, we consider another constraint of dynamics originated from the equation of motion, namely, the Schrödinger equation. When the Hamiltonian is time-independent, the dynamics is represented by a diagonal unitary matrix in the eigenbasis of a given Hamiltonian. Thus, it is worth investigating random unitary matrices diagonal in a fixed basis, which we refer to as *random diagonal-unitary matrices*. When random diagonal-unitary matrices are applied to a fixed state, we obtain phase-random states, an ensemble of states with randomized phases [32]. Phase-random states have a close connection to the study of time-average properties in quantum statistical mechanics.

In this paper, similarly to random unitary matrices, we consider implementations of random diagonal-unitary matrices by quantum circuits and their applications. To this end, we introduce *diagonal-unitary t -designs*. We show that a diagonal-unitary 2-design is exactly achievable by a *phase-random circuit* composed of $O(N^2)$ diagonal gates if we are allowed to use controlled-phase gates with random phases in addition to single qubit phase gates. On the other hand, if we use controlled-Z gates instead of controlled-phase gates, an approximate 2-design is achieved after applying $O(N^2(N + \log 1/\epsilon))$ diagonal gates. These results show that random parameters in genuine two-qubit gates enhance the ability of randomizing phases. We also provide two applications of diagonal-unitary 2-designs implemented by the phase-random circuit. First, we show that, by combining the phase-random circuit with a simple classical procedure requiring $O(N)$ random bits, we can generate an exact 2-design of random states. Our method of generating the 2-design of random states is simpler than previously known implementations, particularly from an experimental point of view. Second, we propose a quantum simulator of canonical states of classical Hamiltonians by using the phase-random circuit where we do not have to pre-

pare the Hamiltonian itself.

This paper is organized as follows. In Sec. II, we give definitions of random unitary matrices, random diagonal-unitary matrices, and corresponding designs, a unitary t -design and a diagonal-unitary t -design. In Sec. III, we present our two main theorems on the convergence of the phase-random circuits and provide the applications. We present proofs of the first theorem and the second theorem in Sec. IV and in Sec. V, respectively. Finally, we summarize and present concluding remarks in Section VI.

II. RANDOM MATRICES AND DESIGNS

We present basic definitions of random (diagonal)-unitary matrices, (phase)-random states and the corresponding designs.

A. Random unitary matrices and random diagonal-unitary matrices

Definition 1 (Random unitary matrices)

Random unitary matrices $\mathcal{U}_{\text{Haar}}$ are the ensemble of unitary matrices uniformly distributed in terms of the Haar measure.

Definition 2 (Random states) Random states Υ_{random} are the ensemble of states $\{U_\mu |\Psi\rangle\}_{U_\mu \in \mathcal{U}_{\text{Haar}}}$ for any fixed state $|\Psi\rangle \in \mathcal{H}$.

Note that the distribution of random states is independent of the choice of the state $|\Psi\rangle$ since the Haar measure is the unitarily invariant.

We also define random diagonal-unitary matrices, which are understood as random unitary matrices with a fixed basis.

Definition 3 (Random diagonal-unitary matrices)

Random diagonal-unitary matrices in an orthonormal basis $\{|u_n\rangle\}$, $\mathcal{U}_{\text{diag}}(\{|u_n\rangle\})$, are an ensemble of diagonal unitary matrices of the form $U_\varphi = \sum_{n=1}^d e^{i\varphi_n} |u_n\rangle\langle u_n|$ where the phases φ_n are uniformly distributed according to the normalized Lebesgue measure $d\varphi = d\varphi_1 \cdots d\varphi_d / (2\pi)^d$ on $[0, 2\pi)^d$.

Definition 4 (Phase-random states) For a given state $|\Psi\rangle = \sum_n r_n e^{i\omega_n} |u_n\rangle$ where $r_n \geq 0$, $\omega_n \in [0, 2\pi)$ for all n , phase-random states $\Upsilon_{\text{phase}}(\{r_n, |u_n\rangle_n\})$ are an ensemble of states $\{U_\varphi |\Psi\rangle\}_{U_\varphi \in \mathcal{U}_{\text{diag}}(\{|u_n\rangle\})}$, which is equivalent to $\{\sum_n e^{i\varphi_n} r_n |u_n\rangle\}_{\varphi_n \in [0, 2\pi)}$.

Random diagonal-unitary matrices are closely related to the study of the typical properties of time evolutions by time-independent Hamiltonians. Under the assumption that the phases of the expansion coefficients in the eigenbasis of a Hamiltonian are randomly distributed after sufficiently long time, investigations of the statistical properties of random diagonal-unitary matrices are

equivalent to the studies of typical properties of such Hamiltonian dynamics. Phase-random states are introduced for this aim and it has been shown that canonical distributions are realized in subsystems when an initial state and the eigenstates of a Hamiltonian satisfy a trade-off relation [22, 32].

B. Designs

In general, a *design* of another ensemble is an ensemble simulating average properties of the original ensemble. A unitary t -design is an ensemble of unitary matrices that simulates up to the t -th order of moments of random unitary matrices on average.

Definition 5 (Unitary t -designs) Let $\mathcal{U}_{\text{Haar}}$ be random unitary matrices. A unitary t -design $\mathcal{U}_{\text{Haar}}^{(t)}$ is a distribution of unitary matrices $\mathcal{U}_{\text{Haar}}^{(t)} = \{U_\mu\}_{d\mu}$ such that

$$\int_{U_\mu \in \mathcal{U}_{\text{Haar}}^{(t)}} U_\mu^{\otimes t} \otimes (U_\mu^\dagger)^{\otimes t} d\mu = \int_{U_\mu \in \mathcal{U}_{\text{Haar}}} U_\mu^{\otimes t} \otimes (U_\mu^\dagger)^{\otimes t} d\mu,$$

where U^\dagger is an Hermite conjugate of U .

Corresponding designs for random states are referred to *spherical t -designs* [33–35], *complex-projective t -designs* [36, 37] or *quantum state t -designs* [29], which are defined as follows.

Definition 6 (Complex projective t -designs)

Let Υ_{random} be random states. A complex projective t -design $\Upsilon_{\text{cp}}^{(t)}$ is a distribution of quantum states $\{|\psi_\mu\rangle\}_{d\mu}$ satisfying

$$\int_{|\psi_\mu\rangle \in \Upsilon_{\text{cp}}^{(t)}} |\psi_\mu\rangle\langle\psi_\mu|^{\otimes t} d\mu = \int_{|\psi_\mu\rangle \in \Upsilon_{\text{random}}} |\psi_\mu\rangle\langle\psi_\mu|^{\otimes t} d\mu.$$

When the distribution of the designs is discrete, the integral is replaced by a summation over a probability distribution, which is sometimes referred to as *weighted t -designs* [38]. By definition, a complex projective t -design cannot be distinguished from random states if we have only t copies of states.

In analogy with unitary t -designs and complex projective t -designs, we define designs for random diagonal-unitary matrices and phase-random states.

Definition 7 (Diagonal-unitary t -designs) Let $\mathcal{U}_{\text{diag}}(\{|u_n\rangle\})$ be random diagonal-unitary matrices in a basis $\{|u_n\rangle\}$. A diagonal-unitary t -design $\mathcal{U}_{\text{diag}}^{(t)}(\{|u_n\rangle\})$ is a distribution of unitary matrices $\{U_\mu\}_{d\mu}$ diagonal in the basis $\{|u_n\rangle\}$, which satisfies that

$$\int_{U_\varphi \in \mathcal{U}_{\text{diag}}^{(t)}} U_\varphi^{\otimes t} \otimes (U_\varphi^\dagger)^{\otimes t} d\varphi = \int_{U_\varphi \in \mathcal{U}_{\text{diag}}} U_\varphi^{\otimes t} \otimes (U_\varphi^\dagger)^{\otimes t} d\varphi.$$

Definition 8 (Toric t -designs) Let

$\Upsilon_{\text{phase}}(\{r_n, |u_n\rangle\})$ be phase-random states. A toric t -design $\Upsilon_{\text{tor}}^{(t)}(\{r_n, |u_n\rangle\})$ is a distribution of states $\{|\psi_\mu\rangle\}_{d\mu}$ satisfying that

$$\int_{|\psi_\varphi\rangle \in \Upsilon_{\text{tor}}^{(t)}} |\psi_\varphi\rangle\langle\psi_\varphi|^{\otimes t} d\varphi = \int_{|\psi_\varphi\rangle \in \Upsilon_{\text{phase}}} |\psi_\varphi\rangle\langle\psi_\varphi|^{\otimes t} d\varphi.$$

Since the parameter space of phase-random states is a d -dimensional torus, we call this type of designs as *toric t -designs*.

The state designs can be generated by applying unitary designs to a state. Inversely, if a set of unitary matrices (diagonal-unitary matrices in the basis $\{|u_n\rangle\}$) generates a complex projective t -design (a toric t -design) for any state $|\Psi\rangle$, then the set is a unitary t -design (a diagonal-unitary t -design).

In this paper, for simplicity, we denote by \mathbb{E} any expectations over a probability distribution. When it is necessary, we explicitly write the space taken over for the expectation, e.g., $\mathbb{E}_{\mathcal{U}_{\text{Haar}}}$ and $\mathbb{E}_{\Upsilon_{\text{cp}}^{(t)}}$. Note that the expectation of unitary matrices is not necessarily unitary and the expectation of states is generally a mixed state.

C. An example of diagonal-unitary t -designs

We show a simple example of a diagonal-unitary t -design where the phases are randomly chosen from a discrete set. This example shows that for achieving diagonal-unitary t -designs, continuous random parameters such as $\alpha \in [0, 2\pi)$ are not necessary and we can use $(t+1)$ -valued discrete random parameters instead.

Proposition 1 A set of unitary matrices $\Omega_t = \{\sum_n e^{i\phi_n} |u_n\rangle\langle u_n|\}_{\{\omega_n\}}$, where ϕ_n is randomly chosen from $\{\frac{2\pi k}{t+1}\}_{k=0,1,\dots,t}$, is a diagonal-unitary t -design in the basis $\{|u_n\rangle\}$.

Proof 1 For $U_\phi = \sum_n e^{i\phi_n} |u_n\rangle\langle u_n|$, $U_\phi^{\otimes t} \otimes U_\phi^{\dagger \otimes t}$ is calculated to

$$U_\phi^{\otimes t} \otimes U_\phi^{\dagger \otimes t} = \sum_{n_1, \dots, n_t=1}^d \sum_{m_1, \dots, m_t=1}^d \exp[i \sum_{k=1}^t (\phi_{n_k} - \phi_{m_k})] |u_{n_1} \dots u_{n_t} u_{m_1} \dots u_{m_t}\rangle \langle u_{n_1} \dots u_{n_t} u_{m_1} \dots u_{m_t}|.$$

For $\mathcal{U}_{\text{diag}}^{(t)}$ and Ω_t , the expectation of an operator X is taken over $\phi_i \in [0, 2\pi)$ and $\phi'_i \in \{\frac{2\pi k}{t+1}\}_{k=0,1,\dots,t}$ for all $i = 1, \dots, d$, respectively, namely,

$$\begin{aligned} \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(t)}} [X] &= \frac{1}{(2\pi)^d} \int_0^{2\pi} X d\phi_1 \dots d\phi_d, \\ \mathbb{E}_{\Omega_t} [X] &= \left(\frac{1}{t+1}\right)^d \sum_{\phi'_1=0, \dots, \frac{2\pi t}{t+1}} \dots \sum_{\phi'_d=0, \dots, \frac{2\pi t}{t+1}} X. \end{aligned}$$

The equation $\mathbb{E}_{\Omega_t} [U^{\otimes t} \otimes U^{\dagger \otimes t}] = \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(t)}} [U^{\otimes t} \otimes U^{\dagger \otimes t}]$ follows from an identity that

$$\begin{aligned} &\frac{1}{(2\pi)^d} \int_0^{2\pi} \exp[i \sum_{k=1}^t (\phi_{n_k} - \phi_{m_k})] d\phi_1 \dots d\phi_d = \\ &\left(\frac{1}{t+1}\right)^d \sum_{\phi'_1=0, \dots, \frac{2\pi t}{t+1}} \dots \sum_{\phi'_d=0, \dots, \frac{2\pi t}{t+1}} \exp[i \sum_{k=1}^t (\phi'_{n_k} - \phi'_{m_k})]. \end{aligned}$$

■

Generation of the matrices of diagonal-unitary t -designs based on Proposition 1 requires global randomizations over the Hilbert space of n qubits. For implementing these matrices by using quantum circuits, we need to decompose diagonal-unitary t -designs into local unitary operations, namely, one- and two-qubit gates, in an efficient way. This is the main concern of this paper.

D. ϵ -approximate t -designs

An ϵ -approximate t -design is an ensemble that approximates the t -design within an error ϵ [8]. For unitary designs and diagonal-unitary designs, we use the diamond norm to evaluate the difference between two matrices. For a superoperator \mathcal{E} on \mathcal{H} , the diamond norm is defined by

$$\|\mathcal{E}\|_\diamond := \sup_d \sup_{X \neq 0} \frac{\|(\mathcal{E} \otimes \text{id}_d)X\|_1}{\|X\|_1},$$

where id_d is an identity operator on another d -dimensional Hilbert space \mathcal{H}' and X is any positive operator on $\mathcal{H} \otimes \mathcal{H}'$ [39]. The diamond norm gives the probability of distinguishing two operations if we are allowed to use an auxiliary system with any dimension.

Definition 9 Let $\mathcal{U}^{(t)}$ be a unitary t -design or a diagonal-unitary t -design. An ϵ -approximate t -designs of $\mathcal{U}^{(t)}$ is a set of unitary operators $\mathcal{U}^{(t, \epsilon)}$ that satisfies

$$\|\mathbb{E}_{\mathcal{U}^{(t)}} [U^{\otimes t} \otimes U^{\dagger \otimes t}] - \mathbb{E}_{\mathcal{U}^{(t, \epsilon)}} [U^{\otimes t} \otimes U^{\dagger \otimes t}]\|_\diamond < \epsilon.$$

For state designs such as complex projective and toric t -designs, we define an ϵ -approximate t -designs in terms of the trace norm.

Definition 10 Let $\Upsilon^{(t)}$ be a complex projective t -design or a toric t -design. An ϵ -approximate t -designs of $\Upsilon^{(t)}$ is a set of states $\Upsilon^{(t, \epsilon)}$ that satisfies

$$\|\mathbb{E}_{\Upsilon^{(t)}} [|\Psi\rangle\langle\Psi|^{\otimes t}] - \mathbb{E}_{\Upsilon^{(t, \epsilon)}} [|\Psi\rangle\langle\Psi|^{\otimes t}]\|_1 < \epsilon,$$

where $\|X\| = \text{Tr}|X|$ is the trace norm.

Since the t -designs are sufficient to perform most of known quantum information tasks using random unitary matrices or random states [40], they have been

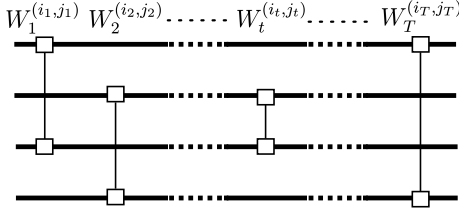


FIG. 1: A phase-random circuit. The vertical line denotes a two-qubit gate $W_t^{(i_t, j_t)}$ randomly selected from a diagonal two-qubit gate set $\mathcal{W}_{\text{diag}}$ on randomly chosen a pair of i_t -th and j_t -th qubits.

intensely studied. Many exact or approximate complex projective t -designs are known for N -qubit systems [28, 35, 37, 41, 43]. The construction of unitary t -designs by using quantum circuits has been also studied [4, 8, 29–31].

Little is known about how to construct random diagonal-unitary t -designs and toric t -designs by quantum circuits. In Ref. [32], it is shown that a quantum circuit composed of controlled-Z and single qubit phase gates generates an ensemble of states of which average amount of entanglement is same as that of phase-random states measured by a linear entropy of a reduced density matrix. However, it had not known whether the circuit achieves a diagonal-unitary t -design or not.

III. MAIN RESULTS AND APPLICATIONS

In this section, we introduce phase-random circuits proposed in [32] and state our main results, i.e., the phase-random circuits can achieve diagonal-unitary 2-designs. Then, we provide two applications of the phase-random circuits.

Note that diagonal-unitary t -designs for $t \geq 2$ can generate large amount of entanglement [32], which implies that diagonal-unitary t -designs for $t \geq 2$ cannot be achieved if we use only one-qubit gates since one-qubit gates cannot increase entanglement. We also note that diagonal-unitary t -designs for general t cannot be achieved by applying random diagonal one- and two-qubit gates since the commutability of gates leads to the lack of parameters. Therefore, we concentrate in the case of $t = 2$.

A. Phase-random circuit

We denote by $|\bar{n}\rangle$ the computational basis where \bar{n} is a binary representation of n ($n = 1, \dots, 2^N$). We investigate achievability of a diagonal-unitary t -design in the computational basis $\mathcal{U}_{\text{diag}}^{(t)}(\{|\bar{n}\rangle\})$ by using *phase-random circuits* presented in the previous work of the authors [32]. For implementations of a diagonal-unitary t -design in a general basis, it is sufficient to apply a uni-

tary operation transforming the computational basis to the basis before and after applying the phase-random circuit.

A phase-random circuit consists of T diagonal two-qubit unitary gates shown in Fig. 1. For the t -th gate, we select two different numbers (i_t, j_t) from $\{1, 2, \dots, N\}$, as well as a two-qubit gate W_t randomly from a given set of diagonal two-qubit gates $\mathcal{W}_{\text{diag}}$. We apply the two-qubit gate W_t on i_t -th and j_t -th qubits. An instance of the circuit is then specified by a set of parameters, $\mathcal{C}_T := \{i_t, j_t, W_t\}_{t=1}^T$, and the unitary operation represented by the circuit is given by $U_T = W_T^{(i_T, j_T)} W_{T-1}^{(i_{T-1}, j_{T-1})} \dots W_1^{(i_1, j_1)}$, where $W_t^{(i_t, j_t)}$ acts on i_t -th and j_t -th qubits. Thus a phase-random circuit consisting of T two-qubit gates is denoted by a set of the unitary operations $\{U_T\}_{\mathcal{C}_T}$.

B. Main results

The main result of this paper is that, if we choose an appropriate two-qubit diagonal gate set $\mathcal{W}_{\text{diag}}$, the phase-random circuit achieves a diagonal-unitary 2-design in the polynomial number of the gates. The necessary number of the gates depends on the choice of the gate set $\mathcal{W}_{\text{diag}}$.

We consider two gate sets. First, we study a gate set given by $\mathcal{W}_{\text{diag}}^{CP} = \{\text{diag}(1, e^{i\alpha}) \otimes \text{diag}(1, e^{i\beta}) \cdot CP(\gamma)\}_{\alpha, \beta, \gamma \in \{0, \frac{2\pi}{3}, \frac{4\pi}{3}\}}$ where $CP(\gamma) = \text{diag}(1, 1, 1, e^{i\gamma})$ is a controlled-phase gate with a phase γ . We refer to the phase-random circuit consisting of this gate set as a *CP phase-random circuit*, where the set of parameters is given by $\mathcal{C}_T^{CP} = \{\alpha_t, \beta_t, \gamma_t\}_{t=1}^T$. Then, we obtain the following theorem:

Theorem 1 The CP phase-random circuit is a diagonal-unitary 2-design in the computational basis if two-qubit gates randomly chosen from a gate set $\mathcal{W}_{\text{diag}}^{CP}$ are applied on all the different pairs of qubits. Thus, the number of the required gates is given by $T = \frac{N(N-1)}{2}$.

Since all gates in the CP phase-random circuit are commutable, they can be applied simultaneously in a practical implementation so that the depth of the circuit is $O(1)$.

Next, we deal with a gate set given by $\mathcal{W}_{\text{diag}}^{CZ} = \{\text{diag}(1, e^{i\alpha}) \otimes \text{diag}(1, e^{i\beta}) \cdot CZ\}_{\alpha, \beta \in \{0, \frac{2\pi}{3}, \frac{4\pi}{3}\}}$, where $CZ := \text{diag}(1, 1, 1, -1)$ is a controlled-Z gate. In this case, we have to choose (i_t, j_t) randomly at each time, so that the set of parameters is given by $\mathcal{C}_T^{CZ} = \{i_t, j_t, \alpha_t, \beta_t\}_{t=1}^T$. We call the corresponding phase-random circuit as *CZ phase-random circuit*. Due to the fact that a controlled-Z gate has no parameter, the circuit converges slowly as stated in Theorem 2.

Theorem 2 The CZ phase-random circuit $\{U_T\}_{\mathcal{C}_T^{CZ}}$ consisting of T two-qubit gates randomly chosen from

a gate set $\mathcal{W}_{\text{diag}}^{CZ}$ applied on a pair of randomly chosen qubits is an ϵ -approximate diagonal-unitary 2-design if $T > T_{\text{conv}}(\epsilon)$ where

$$\begin{aligned} \frac{N^3}{2} \log 2 + \frac{N^2}{2} \log \epsilon^{-1} + O(N^2) &\leq T_{\text{conv}}(\epsilon) \\ &\leq 3N^3 \log 2 + N^2 \log \epsilon^{-1} + O(N^2). \end{aligned}$$

Therefore, the CZ phase-random circuit is an ϵ -approximate diagonal-unitary 2-design after applying $O(N^2(N + \log \epsilon^{-1}))$ two-qubit gates.

In a practical sense, the CZ phase-random circuit has disadvantages comparing to the CP phase-random circuit since it cannot achieve an exact diagonal-unitary 2-design. Moreover, unlike the CP case, the gates in the circuit cannot be applied simultaneously since the dynamics should be *stochastic* by choosing i and j randomly for each gate in order to sufficiently randomize a state, which will be seen in the proof of Theorem 2. However, we present Theorem 2 since the difference between the CZ and the CP phase-random circuits show that the additional parameter of the controlled-phase gates dramatically improve the ability of randomization, which is interesting in the theoretical point of view.

Note that, for both of phase-random circuits, it suffices to take the phases α, β and γ from a discrete set $\{0, \frac{2\pi}{3}, \frac{4\pi}{3}\}$ instead of a continuous set $[0, 2\pi)$ as shown in the example with $t = 2$ of the diagonal-unitary t -designs presented in Subsection II C.

Finally, we emphasize that the phase-random circuits are expected to be easily implemented in experiments since they are composed only of up to two-qubit gates diagonal in the computational basis. Furthermore, in the case of the CP phase-random circuit, all gates can be applied simultaneously, which significantly simplifies the experimental implementation.

C. Applications of phase-random circuits

We show applications of the phase-random circuits. We propose applications for generating a complex projective 2-design and for quantum simulators of canonical states.

1. Generating a complex projective 2-design

The phase-random circuit can generate toric 2-designs exactly for the CP case or approximately for the CZ case. By combining the phase-random circuits with an extra classical procedure, we can also obtain a complex projective 2-design. To show this, we first consider the difference between complex projective and toric 2-designs.

The expectation of states for complex projective t -

designs is calculated to

$$\begin{aligned} \mathbb{E}_{\Upsilon_{\text{random}}} [|\psi\rangle\langle\psi|^{\otimes t}] &:= \int_{|\psi\rangle \in \Upsilon_{\text{random}}} |\psi\rangle\langle\psi|^{\otimes t} d\mu \\ &= \frac{1}{d_{\text{sym}}^{(t)}} \Pi_{\text{sym}}^{(t)}, \end{aligned} \quad (1)$$

where $\Pi_{\text{sym}}^{(t)}$ is a projector onto the symmetric subspace in $\mathcal{H}^{\otimes t}$ and d_{sym} is the dimension of the symmetric subspace. Equation (1) is obtained from Schur's lemma [42].

On the other hand, the expectation of the state for toric t -designs is not a projector $\Pi_{\text{sym}}^{(t)}$. For instance, in N -qubit systems, the expectation of the state for the toric 2-designs $\Upsilon_{\text{tor}}^{(2)}(\{r_n, |u_n\rangle\}_n)$ is calculated to

$$\begin{aligned} \mathbb{E}_{\Upsilon_{\text{tor}}^{(2)}} [|\psi\rangle\langle\psi|^{\otimes 2}] &= \sum_{n=1}^{2^N} r_n^4 |u_n u_n\rangle\langle u_n u_n| \\ &+ 2 \sum_{n>m} r_n^2 r_m^2 \frac{|u_n u_m\rangle + |u_m u_n\rangle}{\sqrt{2}} \frac{\langle u_n u_m| + \langle u_m u_n|}{\sqrt{2}}, \end{aligned}$$

which is not proportional to the projector $\Pi_{\text{sym}}^{(2)}$.

To close the gap, we consider a particular toric 2-design given by $\Upsilon_{\text{tor}}^{(2)}(\{2^{-N/2}, |\bar{n}\rangle\}_n)$ where $\{|\bar{n}\rangle\}$ is the computational basis. The expectation of the state over $\Upsilon_{\text{tor}}^{(2)}(\{2^{-N/2}, |\bar{n}\rangle\}_n)$ is given by

$$\begin{aligned} \mathbb{E}_{\Upsilon_{\text{tor}}^{(2)}} [|\psi\rangle\langle\psi|^{\otimes 2}] &\propto \sum_n |\bar{n}\bar{n}\rangle\langle\bar{n}\bar{n}| \\ &+ 2 \sum_{n>m} \frac{|\bar{n}\bar{m}\rangle + |\bar{m}\bar{n}\rangle}{\sqrt{2}} \frac{\langle\bar{n}\bar{m}| + \langle\bar{m}\bar{n}|}{\sqrt{2}} \\ &= 2\Pi_{\text{sym}}^{(2)} - \sum_n |\bar{n}\bar{n}\rangle\langle\bar{n}\bar{n}|. \end{aligned}$$

This shows that the probabilistic mixture of $\Upsilon_{\text{tor}}^{(2)}(\{2^{-N/2}, |\bar{n}\rangle\}_n)$ and a set of states $\{|\bar{n}\rangle\}$ forms a complex projective 2-design.

The resulting ensemble of states forms a complex projective 2-design since the expectation of the state is proportional to $\Pi_{\text{sym}}^{(2)}$. Thus, the procedure to obtain a complex projective 2-design by using a phase-random circuit is given by

1. With probability $\frac{1}{2^{N+1}}$, choose a random bit \bar{n} and generate a state $|\bar{n}\rangle$.
2. With probability $\frac{2^N}{2^{N+1}}$, perform the CP phase-random circuit with an initial state $|++\dots+\rangle$.

This method of generating the complex projective 2-design requires $O(N)$ random bits and $O(N^2)$ quantum gates in the CP phase-random circuit. The simple structure of the CP phase-random circuit leads to an easier experimental implementation of the design than the previously known results as listed below:

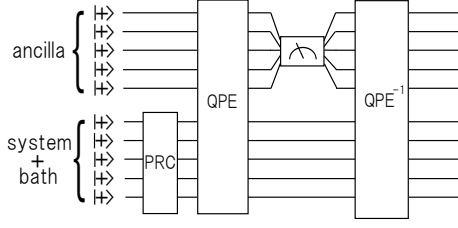


FIG. 2: A quantum circuit to simulate canonical states. The upper half represents ancilla qubits and the lower half represents the system and the thermal bath. The PRC and QPE (QPE⁻¹) are abbreviations of the phase-random circuit and the (inverse of) quantum phase estimation, respectively.

- In Ref. [4], an exact unitary 2-design by using Clifford group is given, which requires $O(N^8)$ bits and $O(N^2)$ quantum gates. In this method, we classically calculate the unitary matrix and decompose it into one- and two-qubit unitary gates. Hence, at each time, it is necessary to repeat the calculation of the gate decomposition and to reconstruct the corresponding quantum circuit.
- In Ref. [8], an ϵ -approximate unitary 2-design is presented where the circuit is composed of one- and two-qubit Clifford gates and some gates are applied probabilistically. The number of gates is $O(N \log 1/\epsilon)$ in their definition of the 2-design corresponding to $O(N(N + \log 1/\epsilon))$ in Definition 9 [29].
- In Ref. [29], an ϵ -approximate unitary 2-design by a random circuit [44] composed of $O(N(N + \log 1/\epsilon))$ gates is proposed. All gates are randomly chosen from a so-called *2-copy gapped* gate set, e.g., a set of controlled-NOT and single qubit rotations.
- In Ref. [31], a local random circuit composed of $O(Nt^4(N + \log 1/\epsilon))$ gates is shown to form an ϵ -approximate unitary t -designs. The circuit is composed of random $SU(4)$ gates acting on nearest neighbor qubits.

In comparison with these results, the main advantage of our method for generating a complex projective 2-design is that it uses only diagonal gates and all gates can be applied simultaneously.

2. Quantum simulator of canonical states

The origin of the thermal equilibration of the system has been studied in the context of random states [19–23]. Consider a system attached with a thermal bath and a Hamiltonian acting on the whole system, $H_{tot} = H_S + H_B + H_{int}$ where H_S , H_B and H_{int} denote the Hamiltonian on the system, that on the thermal bath and an interaction Hamiltonian between them, respectively. Let \mathcal{H}_e be the Hilbert space restricted by a total

energy e . It has been shown that, for almost any states of random states in \mathcal{H}_e , a reduced state in a system is close to the canonical state $\rho_{th} = e^{-\beta H_S} / \text{Tr} e^{-\beta H_S}$ where β is an inverse temperature determined by the total energy e [20]. It is also shown that this is the case for ϵ -approximate complex projective t -designs [40]. Such a typical equilibration is observed even for phase-random states although it depends on $\{r_n, |e_n\rangle\}$ whether the state equilibrates to the canonical state [22, 32]. For instance, in the case of $\Upsilon_{\text{phase}}(\{\frac{1}{\sqrt{d_e}}, |e_i\rangle\})$ where $d_e = \dim \mathcal{H}_e$ and $\{|e_i\rangle\}$ are the eigenstates of H_{tot} in \mathcal{H}_e , the reduced state in the system is almost canonical with high probability.

Besides the fundamental importance on statistical mechanics, these results offer a possibility to simulate canonical states by using random and phase-random states in \mathcal{H}_e , without knowing the eigenstates of the Hamiltonian. Particularly, the phase-random circuit can efficiently simulate canonical states for classical Hamiltonians such as Ising models and classical spin glass models. Since all eigenstates of such Hamiltonians are separable, it is sufficient to prepare phase-random states in the computational basis, which is approximately done by the phase-random circuit.

To see this more clearly, let us consider $N = N_S + N_B$ qubits where N_S and N_B are the number of qubits in the system and that in the thermal bath, respectively. Let $H_{tot} = \sum_i E_i |\bar{n}_i\rangle\langle\bar{n}_i|$ be a classical Hamiltonian where $\{|\bar{n}_i\rangle\}$ is the computational basis. We first define the restricted Hilbert space $\mathcal{H}_e = \text{span}\{|\bar{n}_i\rangle | e - N\delta < E_i < e\}$ where δ is a small number independent of N . The total energy e determines the temperature of the system. For obtaining a canonical state, it is sufficient to generate $\Upsilon_{\text{phase}}(\{\frac{1}{\sqrt{d_e}}, |\bar{n}_i^e\rangle\})$, which is approximately achieved by the quantum circuit presented in Fig. 2.

The circuit is composed of r ancilla qubits initially prepared in $|+\rangle^{\otimes r}$, the phase-random circuit, the quantum phase estimation (QPE) [45, 46] and the projective measurement on the ancilla qubits in the computational basis. The qubits of the system and the bath are also prepared in $|+\rangle^{\otimes N}$, which is the equal-amplitude superposition of all eigenbasis of H_{tot} . At each step, the state changes as follows:

1. The phase-random circuit adds random phases $\{\varphi_k\}$ and generates a state in $\Upsilon_{\text{tot}}^{(2)}(\{2^{-N/2}, |\bar{n}_i\rangle\})$.
2. After QPE, the state is approximately $|\Psi_e\rangle = 2^{-N/2} \sum e^{i\varphi_k} |\bar{n}_k\rangle \otimes |\bar{E}_k\rangle$ where \bar{E}_k is a binary representation of the eigenenergy E_k .
3. By performing the projective measurement $P := \{P_e, P_{\neg e}\}$, where P_e is a projection onto \mathcal{H}_e and $P_{\neg e} = I - P_e$, on the ancilla qubits, the state is probabilistically changed into $\frac{1}{\sqrt{d_e}} \sum_{\bar{n}_k \in \mathcal{H}_e} e^{i\varphi_k} |\bar{n}_k\rangle \otimes |\bar{E}_k\rangle$ if we obtain the outcome corresponding to P_e . Otherwise, the simulation fails.
4. The inverse of QPE changes the state into $|\Psi_f\rangle =$

$(\frac{1}{\sqrt{d_e}} \sum_{\bar{n}_k \in \mathcal{H}_e} e^{i\varphi_k} |\bar{n}_k\rangle) \otimes |+\rangle^r$, so that we obtain a state contained in $\Upsilon_{\text{tot}}^{(2)}(\{\frac{1}{\sqrt{d_e}}, |\bar{n}_i^e\rangle\})$ by tracing out the ancilla qubits.

The error of the simulation mainly originates from QPE for two reasons. First, the eigenenergy is approximated by binary numbers within a precision of 2^{-r} . This approximation results in a round-off error [47]. Second, QPE does not exactly transform the state to $|\Psi_e\rangle$. This is inherited in the final state, resulting in $\sqrt{(1-\epsilon)}|\Psi_f\rangle + \sqrt{\epsilon}|\Psi_{\text{Error}}\rangle$. However, these errors can be suppressed by preparing a large number of ancilla qubits such that $2^{-r} \ll \Delta E$ where ΔE is the minimum energy gap of H_{tot} . For local Hamiltonians, ΔE generally scales exponentially with the system size N , so that $r = \text{poly}(N)$ is sufficient.

Note that obtaining canonical states at low temperatures is generally difficult since the probability to obtain P_e in the projective measurement depends on the temperature as shown in Ref. [23].

Our method is similar to the method proposed in Ref. [23] in the sense that both simulate the *principle of equal weight* on quantum circuits, from which canonical states are derived in statistical mechanics, and exploit the projective measurement to obtain states in \mathcal{H}_e . Contrary to our circuit using only pure states, the circuit in Ref. [23] exploits mixed states and it works for any Hamiltonians. Comparing to the other methods for simulating canonical states [47, 48], our method may be simpler for experiments since all gates except the Fourier transformation in the phase estimation are diagonal in the computational basis and can be applied simultaneously. Thus, the experimental realization of the Fourier transformation immediately leads to the implementation of our circuit. However, our method has a disadvantage that it can efficiently simulate canonical states only of classical Hamiltonians.

D. Sketch of the proofs

In order to prove Theorem 1 and Theorem 2, we analyze how two-qubit diagonal gates in the phase-random circuits transform an initial state $|\phi_0\rangle$. We denote the state after applying T two-qubit diagonal gates by $|\phi_T\rangle = U_T |\phi_0\rangle$. We expand the state $|\phi_T\rangle$ in the Pauli base and investigate the evolution of each coefficient. We denote by \mathbf{p} and \mathbf{q} vectors corresponding to the subscripts of the Pauli base of two N -qubit systems (p_1, \dots, p_N) and (q_1, \dots, q_N) , where $p_i, q_i \in \{0, x, y, z\}$, respectively. Then, the state $|\phi_T\rangle$ is expressed by

$$|\phi_T\rangle\langle\phi_T|^{\otimes 2} = 2^{-N} \sum_{\mathbf{p}, \mathbf{q}} \xi_T(\mathbf{p}, \mathbf{q}) \sigma_{\mathbf{p}} \otimes \sigma_{\mathbf{q}},$$

where $\sigma_{\mathbf{p}} := \sigma_{p_1} \otimes \dots \otimes \sigma_{p_N}$ and $\sigma_{\mathbf{q}} := \sigma_{q_1} \otimes \dots \otimes \sigma_{q_N}$ are tensor products of the Pauli operators. Similarly, we consider a state $|\phi_\varphi\rangle = U_\varphi |\phi_0\rangle$, where U_φ is an element

of a diagonal-unitary 2-design $\mathcal{U}_{\text{diag}}^{(2)}$, and expand $|\phi_\varphi\rangle$ in the Pauli base

$$|\phi_\varphi\rangle\langle\phi_\varphi|^{\otimes 2} = 2^{-N} \sum_{\mathbf{p}, \mathbf{q}} \xi_\varphi(\mathbf{p}, \mathbf{q}) \sigma_{\mathbf{p}} \otimes \sigma_{\mathbf{q}}.$$

In terms of the expansion coefficients $\xi_T(\mathbf{p}, \mathbf{q})$ and $\xi_\varphi(\mathbf{p}, \mathbf{q})$, our goal is to show that for any initial state $|\phi_0\rangle$,

$$\forall(\mathbf{p}, \mathbf{q}), \left| \mathbb{E}_{\mathcal{C}_T} [\xi_T(\mathbf{p}, \mathbf{q})] - \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})] \right| < \frac{\epsilon}{2^{2N}}, \quad (2)$$

after sufficiently large T . By applying the similar argument presented in Ref. [29], we can show that if Eq. (2) holds for any initial state, the phase-random circuit is an ϵ -approximate diagonal-unitary 2-design.

In Section IV, we present that the CP phase-random circuit achieves Eq. (2) for $\epsilon = 0$ after applying $O(N^2)$ two-qubit diagonal gates, which gives the proof of Theorem 1. In Section V, we show that Eq. (2) holds for the CZ phase-random circuit after applying $O(N^2(N + \log 1/\epsilon))$ two-qubit diagonal gates, implying Theorem 2.

IV. CP PHASE-RANDOM CIRCUIT

We present a proof of Theorem 1. To prove Theorem 1, we follow the transformation of the expectation of an expansion coefficient $\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})]$ by the CP phase-random circuit.

A. Notations

We introduce several notations that simplify our investigations. Since the way how $\xi_T(\mathbf{p}, \mathbf{q})$ is transformed depends on $\sigma_{\mathbf{p}}$ and $\sigma_{\mathbf{q}}$, it is convenient to define subsets in $\{1, \dots, N\}$ that specify the locations of σ_w ($w = 0, x, y, z$) in $\sigma_{\mathbf{p}}$ and $\sigma_{\mathbf{q}}$;

$$\begin{aligned} \Gamma^{(+)}(\mathbf{p}, \mathbf{q}) &:= \{i \in \{1, \dots, N\} | p_i = q_i \in \{x, y\}\}, \\ \Gamma^{(-)}(\mathbf{p}, \mathbf{q}) &:= \{i \in \{1, \dots, N\} | p_i = \bar{q}_i \in \{x, y\}\}, \\ \Lambda^{(+)}(\mathbf{p}, \mathbf{q}) &:= \{i \in \{1, \dots, N\} | p_i = q_i \in \{0, z\}\}, \\ \Lambda^{(-)}(\mathbf{p}, \mathbf{q}) &:= \{i \in \{1, \dots, N\} | p_i = \bar{q}_i \in \{0, z\}\}, \end{aligned}$$

where the bar sign of \bar{q}_i represents to take a self-inverse ‘flip’ map defined by $\bar{0} = z$ and $\bar{x} = y$. We denote the number of elements in each subset by the corresponding small letters, e.g., $\gamma^{(\pm)}(\mathbf{p}, \mathbf{q})$ ($\lambda^{(\pm)}(\mathbf{p}, \mathbf{q})$) is the number of elements in $\Gamma^{(\pm)}(\mathbf{p}, \mathbf{q})$ ($\Lambda^{(\pm)}(\mathbf{p}, \mathbf{q})$). We also denote the union of $\Gamma^{(+)}(\mathbf{p}, \mathbf{q})$ and $\Gamma^{(-)}(\mathbf{p}, \mathbf{q})$ ($\Lambda^{(+)}(\mathbf{p}, \mathbf{q})$ and $\Lambda^{(-)}(\mathbf{p}, \mathbf{q})$) by $\Gamma(\mathbf{p}, \mathbf{q})$ ($\Lambda(\mathbf{p}, \mathbf{q})$). Similarly, the number of elements in $\Gamma(\mathbf{p}, \mathbf{q})$ ($\Lambda(\mathbf{p}, \mathbf{q})$) is denoted by $\gamma(\mathbf{p}, \mathbf{q})$ ($\lambda(\mathbf{p}, \mathbf{q})$). By definition, $\gamma(\mathbf{p}, \mathbf{q}) = \gamma^{(+)}(\mathbf{p}, \mathbf{q}) + \gamma^{(-)}(\mathbf{p}, \mathbf{q})$ and $\lambda(\mathbf{p}, \mathbf{q}) = \lambda^{(+)}(\mathbf{p}, \mathbf{q}) + \lambda^{(-)}(\mathbf{p}, \mathbf{q})$. For simplicity, we often omit the part (\mathbf{p}, \mathbf{q}) in equations when there is no ambiguity.

We also define a function $f_S(\mathbf{p})$ of \mathbf{p} , where $S = \{i_1, \dots, i_s\}$ is a subset of $\{1, \dots, N\}$, such as $f_S(\mathbf{p}) = (p_1, \dots, \bar{p}_{i_1}, \dots, \bar{p}_{i_s}, \dots, p_N)$. That is, the function f_S flips all elements of \mathbf{p} in the set S . For instance, $f_{1,3}(y, x, 0) = (\bar{y}, x, \bar{0}) = (x, x, z)$. We also denote $(f_S(\mathbf{p}), f_S(\mathbf{q}))$ simply by $f_S(\mathbf{p}, \mathbf{q})$.

B. Calculation of $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]$

We calculate $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]$ for an initial state $|\phi_0\rangle = \sum r_n e^{i\omega_n} |\bar{n}\rangle$ ($\forall n, r_n \geq 0, \omega_n \in [0, 2\pi)$). Since $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})] = 2^{-N} \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\langle \phi_\varphi | \sigma_{\mathbf{p}} | \phi_\varphi \rangle \langle \phi_\varphi | \sigma_{\mathbf{q}} | \phi_\varphi \rangle]$, some calculations lead to the coefficients. For (\mathbf{p}, \mathbf{q}) satisfying $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) = N$, we obtain

$$\begin{aligned} \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})] &= 2^{-N} \sum_{n,m} r_n^2 r_m^2 \prod_{j \in \Lambda^{(+)}} \delta_{n_j m_j} \\ &\times \prod_{j \in \Lambda^{(-)}} \delta_{n_j m_j} (1 - 2n_j) \\ &\times \prod_{j \in \Gamma^{(+)}} (1 - \delta_{n_j m_j}) \\ &\times \prod_{j \in \Gamma^{(-)}} -i(1 - \delta_{n_j m_j})(\delta_{p_j x} - \delta_{p_j y})(1 - 2n_j), \quad (3) \end{aligned}$$

where δ_{nm} is Kronecker delta and $n_1 n_2 \dots n_N$ ($m_1 m_2 \dots m_N$) is a binary representation of $n - 1$

$(m - 1)$. In other cases, $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})] = 0$.

C. Transformation of $\mathbb{E}_{\mathcal{C}_T^{CP}} [\xi_T(\mathbf{p}, \mathbf{q})]$

We consider the transformation of a state by the CP phase-random circuit. Let us consider the expansion of $\mathbb{E}_{\mathcal{C}_T^{CP}} [|\phi_T\rangle \langle \phi_T|^{\otimes 2}]$ in the Pauli base given by

$$\mathbb{E}_{\mathcal{C}_T^{CP}} [|\phi_T\rangle \langle \phi_T|^{\otimes 2}] = 2^{-N} \sum_{\mathbf{p}, \mathbf{q}} \mathbb{E}_{\mathcal{C}_T^{CP}} [\xi_T(\mathbf{p}, \mathbf{q})] \sigma_{\mathbf{p}} \otimes \sigma_{\mathbf{q}}.$$

For simplicity, hereafter we omit the subscript \mathcal{C}_T^{CP} for the expectation values. By applying $W_{T+1}(\alpha_{T+1}, \beta_{T+1}, \gamma_{T+1})$ on a pair of qubits specified by two numbers (i, j) , the expectation of coefficients changes to

$$\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})] = 2^{-2N} \sum_{(\mathbf{p}', \mathbf{q}')} G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') \mathbb{E}[\xi_{T+1}(\mathbf{p}', \mathbf{q}')],$$

where the matrix $G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$ is given by

$$G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') = \mathbb{E}[\text{Tr} \sigma_{\mathbf{p}} W_{T+1} \sigma_{\mathbf{p}'} W_{T+1}^\dagger \text{Tr} \sigma_{\mathbf{q}} W_{T+1} \sigma_{\mathbf{q}'} W_{T+1}^\dagger].$$

In Appendix A, we present a derivation of the matrix $G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$. Then, we obtain

$$\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})] = \begin{cases} \mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] & \text{if } i, j \in \Lambda(\mathbf{p}, \mathbf{q}), \\ \frac{1}{4} (\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] \pm \mathbb{E}[\xi_T(f_i(\mathbf{p}, \mathbf{q}))] \\ \quad + \mathbb{E}[\xi_T(f_j(\mathbf{p}, \mathbf{q}))] \pm \mathbb{E}[\xi_T(f_{ij}(\mathbf{p}, \mathbf{q}))]) & \text{if } i \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q}), j \in \Lambda(\mathbf{p}, \mathbf{q}), \\ \frac{1}{4} (\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] + v \mathbb{E}[\xi_T(f_j(\mathbf{p}, \mathbf{q}))] \\ \quad + u \mathbb{E}[\xi_T(f_i(\mathbf{p}, \mathbf{q}))] + uv \mathbb{E}[\xi_T(f_{ij}(\mathbf{p}, \mathbf{q}))]) & \text{if } i \in \Gamma^{(u)}(\mathbf{p}, \mathbf{q}), j \in \Gamma^{(v)}(\mathbf{p}, \mathbf{q}), \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where u and v is ± 1 .

Note that a set of $\Lambda(\mathbf{p}, \mathbf{q})$ and a set of $\Gamma(\mathbf{p}, \mathbf{q})$ are both invariant under the transformations since the transformations are composed of a function $f_S(\mathbf{p}, \mathbf{q})$ which flips x (y) to y (x) and 0 (z) to z (0).

From the last case of Eq. (4), for any pairs of indices (\mathbf{p}, \mathbf{q}) satisfying $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) < N$, we have $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] = 0$ for any T once after $i \notin \Lambda(\mathbf{p}, \mathbf{q}) \cup \Gamma(\mathbf{p}, \mathbf{q})$ is chosen. The other cases show that when one of (i, j) is in $\Gamma(\mathbf{p}, \mathbf{q})$, we take the uniform average of the ‘flipped’ terms. In particular, when $i \in \Gamma^{(-)}(\mathbf{p}, \mathbf{q})$ is selected, the term acquire a negative sign. Hence, after all pairs of qubits are drawn, the expectation of the state becomes an average of all flipped terms with ap-

propriate negative signs. To express the expectation, it is convenient to introduce $\Gamma_{\text{even}}^{(-)}(\mathbf{p}, \mathbf{q})$ and $\Gamma_{\text{odd}}^{(-)}(\mathbf{p}, \mathbf{q})$ as subsets of $\Gamma^{(-)}(\mathbf{p}, \mathbf{q})$ of which the number of elements is even and odd, respectively. We also introduce $S_{\text{even}}(\mathbf{p}, \mathbf{q})$ and $S_{\text{odd}}(\mathbf{p}, \mathbf{q})$, which are defined by $S_{\text{even}}(\mathbf{p}, \mathbf{q}) = \{f_s(\mathbf{p}, \mathbf{q}) \text{ where } s \subset \Lambda(\mathbf{p}, \mathbf{q}) \cup \Gamma^{(+)}(\mathbf{p}, \mathbf{q}) \cup \Gamma_{\text{even}}^{(-)}(\mathbf{p}, \mathbf{q})\}$ and $S_{\text{odd}}(\mathbf{p}, \mathbf{q}) = \{f_s(\mathbf{p}, \mathbf{q}) \text{ where } s \subset \Lambda(\mathbf{p}, \mathbf{q}) \cup \Gamma^{(+)}(\mathbf{p}, \mathbf{q}) \cup \Gamma_{\text{odd}}^{(-)}(\mathbf{p}, \mathbf{q})\}$. Then, we obtain the following proposition.

Proposition 2 After applying W_{ij} to all combinations of i and j , which requires $T_{CP} = \frac{N(N-1)}{2}$ two-qubit gates,

the coefficient converges to

$$\mathbb{E}[\xi_{TCP}(\mathbf{p}, \mathbf{q})] = 2^{-N} \left[\sum_{S_{even}(\mathbf{p}, \mathbf{q})} - \sum_{S_{odd}(\mathbf{p}, \mathbf{q})} \right] \xi_0(\mathbf{p}', \mathbf{q}'), \quad (5)$$

where the summation is taken over all $(\mathbf{p}', \mathbf{q}') \in S_{even(odd)}(\mathbf{p}, \mathbf{q})$.

Finally, by calculating Eq. (5) explicitly, we obtain that for any initial states and for all (\mathbf{p}, \mathbf{q}) ,

$$\mathbb{E}[\xi_{TCP}(\mathbf{p}, \mathbf{q})] = \mathbb{E}_{\mathcal{U}_{diag}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})].$$

For the detailed calculation, see Appendix B. This concludes the proof of Theorem 1.

V. CZ PHASE-RANDOM CIRCUIT

Theorem 2 is shown by proving the following two lemmas. The first lemma guarantees that the state transformed by the CZ phase-random circuit for any initial state converges to the corresponding toric 2-design.

Lemma 1 For any initial state and $\forall(\mathbf{p}, \mathbf{q})$,

$$\lim_{T \rightarrow \infty} \left| \mathbb{E}_{CZ} [\xi_T(\mathbf{p}, \mathbf{q})] - \mathbb{E}_{\mathcal{U}_{diag}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})] \right| = 0.$$

$$\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})] = \begin{cases} \mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] & \text{if } i, j \in \Lambda(\mathbf{p}, \mathbf{q}), \\ \frac{1}{2}(\mathbb{E}[\xi_T(f_j(\mathbf{p}, \mathbf{q}))] \pm \mathbb{E}[\xi_T(f_{ij}(\mathbf{p}, \mathbf{q}))]) & \text{if } i \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q}), j \in \Lambda(\mathbf{p}, \mathbf{q}), \\ \frac{1}{4}(\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] + v\mathbb{E}[\xi_T(f_j(\mathbf{p}, \mathbf{q}))] \\ \quad + u\mathbb{E}[\xi_T(f_i(\mathbf{p}, \mathbf{q}))] + uv\mathbb{E}[\xi_T(f_{ij}(\mathbf{p}, \mathbf{q}))]) & \text{if } i \in \Gamma^{(u)}(\mathbf{p}, \mathbf{q}), j \in \Gamma^{(v)}(\mathbf{p}, \mathbf{q}), \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

where $u, v \in \{+, -\}$ (see Appendix A). The transformation is different from that of the CP phase-random circuits given by Eq. (4) only when $i \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q})$ and $j \in \Lambda(\mathbf{p}, \mathbf{q})$. This prevents the circuit from efficiently randomizing the corresponding phases. Thus, we have to introduce stochastic transformations by choosing (i, j) randomly for achieving a diagonal-unitary 2-design.

2. Modified phase-random circuit

Similarly to the CP case, for any pairs of indices (\mathbf{p}, \mathbf{q}) satisfying $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) < N$, we have $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] = 0$ for any T once after $i \notin \Lambda(\mathbf{p}, \mathbf{q}) \cup \Gamma(\mathbf{p}, \mathbf{q})$ is chosen. In order to avoid the complication by dealing with such (\mathbf{p}, \mathbf{q}) , we first apply a two-qubit gate $W_{i,j}(\alpha, \beta)$ on all neighboring qubits $(2k-1, 2k)$. We denote this unitary operations

The second lemma states that the convergence time $T_{conv}(\epsilon)$ defined in Theorem 2 scales cubic of the system size N .

Lemma 2 For any initial state, the convergence time $T_{conv}(\epsilon)$ satisfies

$$\begin{aligned} \frac{N^3}{2} \log 2 + \frac{N^2}{2} \log \epsilon^{-1} + O(N^2) &\leq T_{conv}(\epsilon) \\ &\leq 3N^3 \log 2 + N^2 \log \epsilon^{-1} + O(N^2). \end{aligned}$$

Therefore, $T_{conv}(\epsilon) = O(N^3) + O(N^2) \log \epsilon^{-1}$.

We present the proof of Lemma 1 in Subsection V A and the proof of Lemma 2 in Subsection V B.

A. Convergence of the distribution

1. Transformation of $\mathbb{E}_{CZ} [\xi_T(\mathbf{p}, \mathbf{q})]$

We consider the transformation of a state by the CZ phase-random circuit. Similarly to the CP case, we obtain

by $\tilde{W} = W_{N,N-1} W_{N-2,N-3} \cdots W_{2,1}$. (see Fig. 3) When N is odd, we define \tilde{W} by $W_{N,N-1} W_{N-1,N-2} \cdots W_{2,1}$. The number of the two-qubit gates required to perform \tilde{W} is $T_{\tilde{W}} = \lceil N/2 \rceil$ where $\lceil n \rceil$ is the smallest integer larger or equal to n . Note that \tilde{W} is composed of commutable gates and is deterministic. Hence, they can be applied simultaneously.

In an analogous way with Proposition 2, we obtain the following proposition.

Proposition 3 If $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) < N$, $\mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})] = 0$. If (\mathbf{p}, \mathbf{q}) satisfies $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) = N$,

$$\begin{aligned} \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})] &= 2^{-\gamma} \left[\sum_{s \in \Gamma_{even}^{(-)}} - \sum_{s \in \Gamma_{odd}^{(-)}} \right] \sum_{s' \in \Gamma^{(+)}} \xi_0(f_{s \cup s'} \circ f_{\tilde{\Lambda}}(\mathbf{p}, \mathbf{q})), \end{aligned}$$

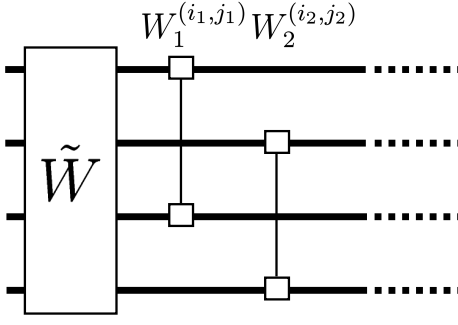


FIG. 3: A modified phase-random circuit. A N -qubit unitary operation \tilde{W} is applied in advance of the phase-random circuit. The unitary operation \tilde{W} is composed of two-qubits diagonal gates acting on neighboring qubits, $\tilde{W} = W_{N,N-1} \cdots W_{3,2} W_{2,1}$.

where (\mathbf{p}, \mathbf{q}) is omitted for simplicity and $\tilde{\Lambda}(\mathbf{p}, \mathbf{q})$ is a set of $i \in \Lambda(\mathbf{p}, \mathbf{q})$ such that i is paired with an element of $\Gamma(\mathbf{p}, \mathbf{q})$ in \tilde{W} .

3. Transformation of $\mathbb{E}[\xi_t(\mathbf{p}, \mathbf{q})]$ after applying \tilde{W}

By using the modified phase-random circuit, after applying \tilde{W} that consists of $T_{\tilde{W}}$ two-qubit gates, additional T' two-qubit diagonal gates randomly selected from $\mathcal{W}_{\text{diag}}^{CP}$ are applied on a randomly chosen pair of qubits. The following proposition shows how $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ relates to $\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})]$ for $T = T_{\tilde{W}} + T'$ where $T' > 0$.

Proposition 4 For $T > T_{\tilde{W}}$,

$$\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})] = \sum_{(\mathbf{p}', \mathbf{q}')} \mathcal{G}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') \mathbb{E}[\xi_T(\mathbf{p}', \mathbf{q}')], \quad (7)$$

where $\mathcal{G}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$ is equal to $\frac{\lambda(\lambda-1) + \gamma(\gamma-1)}{N(N-1)}$ if $(\mathbf{p}', \mathbf{q}') = (\mathbf{p}, \mathbf{q})$, $\frac{2\gamma}{N(N-1)}$ if $(\mathbf{p}', \mathbf{q}') = f_i(\mathbf{p}, \mathbf{q})$ for $i \in \Lambda(\mathbf{p}, \mathbf{q})$ and 0 otherwise.

Proof 2 The equation (7) is equal to

$$\mathbb{E}[\xi_{T+1}(\mathbf{p}, \mathbf{q})] = \frac{\lambda(\lambda-1) + \gamma(\gamma-1)}{N(N-1)} \mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] + \frac{2\gamma}{N(N-1)} \sum_{i \in \Lambda(\mathbf{p}, \mathbf{q})} \mathbb{E}[\xi_T(f_i(\mathbf{p}, \mathbf{q}))]. \quad (8)$$

We show Eq. (8) recursively.

Firstly, we show that $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})] = \mathbb{E}[\xi_{T_{\tilde{W}}}(f_i(\mathbf{p}, \mathbf{q}))]$ when $i \in \Lambda(\mathbf{p}, \mathbf{q})$ and $j \in \Gamma(\mathbf{p}, \mathbf{q})$, and $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})] = \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]$ otherwise. This is derived by using Eq. (6) and Proposition 3 in the following way. When $i, j \in \Lambda(\mathbf{p}, \mathbf{q})$, Eq. (6) implies $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})] = \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]$. When $i \in \Lambda(\mathbf{p}, \mathbf{q})$ and

$j \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q})$, $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})]$ is calculated to be

$$\begin{aligned} \mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})] &= \frac{1}{2} \mathbb{E}[\xi_{T_{\tilde{W}}}(f_i(\mathbf{p}, \mathbf{q}))] \pm \frac{1}{2} \mathbb{E}[\xi_{T_{\tilde{W}}}(f_{ij}(\mathbf{p}, \mathbf{q}))] \\ &= \mathbb{E}[\xi_{T_{\tilde{W}}}(f_i(\mathbf{p}, \mathbf{q}))], \end{aligned}$$

where the second line of the equation is obtained from a property of $\mathbb{E}[\xi_{T_{\tilde{W}}}(f_i(\mathbf{p}, \mathbf{q}))]$ such that

$$\forall j \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q}), \mathbb{E}[\xi_{T_{\tilde{W}}}(f_j(\mathbf{p}, \mathbf{q}))] = \pm \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]. \quad (9)$$

When $i, j \in \Gamma(\mathbf{p}, \mathbf{q})$, a direct calculation shows that $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})] = \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]$ by using Eq. (9). Thus, we obtain the statement.

Since the probability that $j \in \Gamma(\mathbf{p}, \mathbf{q})$ is chosen for a fixed $i \in \Lambda(\mathbf{p}, \mathbf{q})$ is given by $\gamma/\binom{N}{2}$, where $\binom{N}{2}$ is a binomial coefficient, Eq. (8) is shown for $T = T_{\tilde{W}}$. Moreover, $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})]$ also satisfies $\mathbb{E}[\xi_{T_{\tilde{W}}+1}(f_i(\mathbf{p}, \mathbf{q}))] = \pm \mathbb{E}[\xi_{T_{\tilde{W}}+1}(\mathbf{p}, \mathbf{q})]$ for $i \in \Gamma^{(\pm)}(\mathbf{p}, \mathbf{q})$, so that Eq. (8) is recursively obtained. ■

Proposition 4 implies that, for $T > T_{\tilde{W}}$, $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ is given by a convex sum of $\mathbb{E}[\xi_{T_{\tilde{W}}}(f_s(\mathbf{p}, \mathbf{q}))]$ where s is a subset of $\Lambda(\mathbf{p}, \mathbf{q})$. Hence, if we specify $L^{(\pm)} \subsetneq \{1, \dots, N\}$ where $L^{(+)} \cap L^{(-)} = \emptyset$, the transformation of $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ is closed in $\Sigma(L^{(+)}, L^{(-)}) := \{(\mathbf{p}, \mathbf{q}) | \Lambda^{(\pm)}(\mathbf{p}, \mathbf{q}) = L^{(\pm)}\}$. In Proposition 5, we consider the transformation in $\Sigma(L^{(+)}, L^{(-)})$ and derive the stationary distribution $\mathbb{E}[\xi_{\infty}(\mathbf{p}, \mathbf{q})] := \lim_{T \rightarrow \infty} \mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ for $(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})$.

In order to obtain Proposition 5, we use the Perron-Frobenius theorem [49] for *irreducible* and *aperiodic* non-negative matrices M . Irreducibility is a property such that, for all i and j there exists a natural number n such that $(M^n)_{ij} > 0$ and aperiodicity is a property that $M_{ii} > 0$ for all i . A non-negative matrix implies that for $M_{ij} \geq 0$ for all i and j . The Perron-Frobenius theorem is given by the following statement.

Theorem 3 (Perron-Frobenius theorem [49]) If a non-negative matrix M is irreducible and aperiodic, the maximum eigenvalue $\lambda > 0$ is uniquely determined. Let $|\lambda\rangle$ be the corresponding eigenvector to the maximum eigenvalue, then, $\lim_{n \rightarrow \infty} (\frac{1}{\lambda} M)^n = |\lambda\rangle \langle \lambda|$.

In addition to the irreducibility and the aperiodicity, when a non-negative matrix M is *bistochastic*, that is, $\sum_i M_{ij} = \sum_j M_{ij} = 1$, it is known that the maximum eigenvalue λ is equal to 1. By applying these facts, we obtain Proposition 5.

Proposition 5 Let $L^{(\pm)}$ be a proper subset of $\{1, \dots, N\}$ satisfying $L^{(+)} \cap L^{(-)} = \emptyset$ and $L := L^{(+)} \cup L^{(-)} \neq \{1, \dots, N\}$. For $(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})$, the stationary distribution $\mathbb{E}[\xi_{\infty}(\mathbf{p}, \mathbf{q})]$ is uniform in $\Sigma(L^{(+)}, L^{(-)})$, that is,

$$\mathbb{E}[\xi_{\infty}(\mathbf{p}, \mathbf{q})] = \frac{1}{2^l} \sum_{(\mathbf{p}', \mathbf{q}') \in \Sigma(L^{(+)}, L^{(-)})} \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}', \mathbf{q}')], \quad (10)$$

where l is the number of elements of L which is also equal to $\lambda(\mathbf{p}, \mathbf{q})$. Moreover, for any (\mathbf{p}, \mathbf{q}) , we obtain

$$\mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})] = \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]. \quad (11)$$

Proof 3 In the case of $L = \{1, \dots, N\}$, it is straightforward to show Eq. (11) from Eq. (6). When $L \neq \{1, \dots, N\}$, we can obtain Eq. (10) by applying the Perron-Frobenius theorem to the matrix \mathcal{G} in $\Sigma(L^{(+)}, L^{(-)})$. If we restrict the matrix \mathcal{G} in $\Sigma(L^{(+)}, L^{(-)})$, it is straightforward to see that \mathcal{G} is an irreducible, aperiodic and bistochastic non-negative matrix in $\Sigma(L^{(+)}, L^{(-)})$. Hence, the Perron-Frobenius theorem guarantees that there exists a unique stationary distribution in $\Sigma(L^{(+)}, L^{(-)})$. Since the evolution governed by \mathcal{G} in $\Sigma(L^{(+)}, L^{(-)})$ is uniform, the stationary distribution is also uniform, leading that, $\forall(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})$,

$$\mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})] = \frac{1}{2^l} \sum_{(\mathbf{p}', \mathbf{q}') \in \Sigma(L^{(+)}, L^{(-)})} \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}', \mathbf{q}')].$$

Thus, we obtain Eq. (10).

Next, we show Eq. (11). For (\mathbf{p}, \mathbf{q}) satisfying $\lambda(\mathbf{p}, \mathbf{q}) + \gamma(\mathbf{p}, \mathbf{q}) < N$, $\mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})] = 0$ since all $\mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]$ in the right hand side of Eq. (10) are zero as shown in Proposition 3. Since Eq. (3) implies $\mathbb{E}[\xi_{\mathcal{U}_{\text{diag}}^{(2)}}(\mathbf{p}, \mathbf{q})] = 0$ for such (\mathbf{p}, \mathbf{q}) , we obtain $\mathbb{E}[\xi_{T_\infty}(\mathbf{p}, \mathbf{q})] = \mathbb{E}[\xi_{\mathcal{U}_{\text{diag}}^{(2)}}(\mathbf{p}, \mathbf{q})]$.

When (\mathbf{p}, \mathbf{q}) satisfies $\lambda(\mathbf{p}, \mathbf{q}) + \gamma(\mathbf{p}, \mathbf{q}) = N$, we substitute $\mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})]$ given in Proposition 3 into Eq. (10), and obtain

$$\mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})] = 2^{-N} \left[\sum_{S_{\text{even}}(\mathbf{p}, \mathbf{q})} - \sum_{S_{\text{odd}}(\mathbf{p}, \mathbf{q})} \right] \xi_0(\mathbf{p}', \mathbf{q}').$$

As shown in Appendix B, this is equal to $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]$. ■

B. Convergence time for the phase-random circuits

In this section, we investigate the convergence time $T_{\text{conv}}(\epsilon)$ defined by $\forall T > T_{\text{conv}}(\epsilon)$,

$$\|\mathbb{E}_{C_T^{\text{CZ}}} [U_T^{\otimes 2} \otimes U_T^{\dagger \otimes 2}] - \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [U_\varphi^{\otimes 2} \otimes U_\varphi^{\dagger \otimes 2}]\|_\diamond < \epsilon. \quad (12)$$

As explained in Section IIID, a sufficient condition for Eq. (12) to hold is given by

$$\forall(\mathbf{p}, \mathbf{q}), \left| \mathbb{E}_{C_T^{\text{CZ}}} [\xi_T(\mathbf{p}, \mathbf{q})] - \mathbb{E}_{C_\infty^{\text{CZ}}} [\xi_\infty(\mathbf{p}, \mathbf{q})] \right| < \frac{\epsilon}{2^{2N}}. \quad (13)$$

Note that $\mathbb{E}_{C_\infty^{\text{CZ}}} [\xi_\infty(\mathbf{p}, \mathbf{q})] = \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]$ from Lemma 1. Similarly, we can obtain a necessary condition for Eq. (12) to be satisfied by evaluating a lower bound of the diamond norm:

$$\forall(\mathbf{p}, \mathbf{q}), \left| \mathbb{E}_{C_T^{\text{CZ}}} [\xi_T(\mathbf{p}, \mathbf{q})] - \mathbb{E}_{C_\infty^{\text{CZ}}} [\xi_\infty(\mathbf{p}, \mathbf{q})] \right| < \epsilon. \quad (14)$$

We derive an upper and a lower bound of $T_{\text{conv}}(\epsilon)$ by using Eqs. (13) and (14), respectively, and prove the Lemma 2 stating that for any initial state,

$$\begin{aligned} \frac{N^3}{4} \log 2 + \frac{N^2}{4} \log \epsilon^{-1} + O(N^2) &\leq T_{\text{conv}}(\epsilon) \\ &\leq 3N^3 \log 2 + N^2 \log \epsilon^{-1} + O(N^2). \end{aligned}$$

Note that, for (\mathbf{p}, \mathbf{q}) satisfying $\lambda(\mathbf{p}, \mathbf{q}) + \gamma(\mathbf{p}, \mathbf{q}) < N$, $T_{\text{conv}}(\epsilon) \leq T_{\tilde{W}} = \lceil N/2 \rceil$ since $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] = 0$ for $T \geq T_{\tilde{W}}$, and, for (\mathbf{p}, \mathbf{q}) satisfying $\lambda(\mathbf{p}, \mathbf{q}) = N$, $T_{\text{conv}}(\epsilon) = 0$ since $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})] = \xi_0(\mathbf{p}, \mathbf{q})$. In the following, we consider only (\mathbf{p}, \mathbf{q}) such that $\gamma(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p}, \mathbf{q}) = N$ and $\lambda(\mathbf{p}, \mathbf{q}) \neq N$.

In order to show Lemma 2, we use techniques of the Markov chains [50]. We provide a brief introduction of the Markov chains in Appendix C. We map the transformation of $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ into a Markov chain and give a lower and an upper bounds of the convergence time.

1. Map to a Markov chain

We present a map from the transformation of $\mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})]$ by the (modified) CZ phase-random circuit to a Markov chain. Since the transformation in the phase-random circuit is deterministic for $T \leq T_{\tilde{W}}$, we consider the evolution for $T > T_{\tilde{W}}$ as a Markov chain. As shown in Proposition 4, the transformation of $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ is closed in $\Sigma(L^{(+)}, L^{(-)})$ and $\mathcal{G}[(\mathbf{p}, \mathbf{q}); (\mathbf{p}', \mathbf{q}')] satisfies a Markovian property. Moreover, it is observed from Proposition 4 that $\mathcal{G}[(\mathbf{p}, \mathbf{q}); (\mathbf{p}', \mathbf{q}')] is equivalent to a transition matrix of a random walk on a l -dimensional hypercube where each vertex is given by $(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})$. Note that \mathcal{G} is irreducible and aperiodic in $\Sigma(L^{(+)}, L^{(-)})$. However, $\mathbb{E}[\xi_T(\mathbf{p}, \mathbf{q})]$ cannot be regarded as a probability distribution of a Markov chain since they are not necessarily to be non-negative. Instead, we define a probability distribution in the following way.$$

We set the initial probability distribution $\{\Pi_0(\mathbf{p}, \mathbf{q})\}$ on the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$ to be

$$\Pi_0(\mathbf{p}, \mathbf{q}) := \frac{\mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})] - \Pi_{\min}(L^{(+)}, L^{(-)})}{\Pi_{\text{sum}}(L^{(+)}, L^{(-)})},$$

where

$$\Pi_{\min}(L^{(+)}, L^{(-)}) := \min_{(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})} \mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})],$$

and

$$\begin{aligned} \Pi_{\text{sum}}(L^{(+)}, L^{(-)}) &:= \\ &\sum_{(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})} \left(\mathbb{E}[\xi_{T_{\tilde{W}}}(\mathbf{p}, \mathbf{q})] - \Pi_{\min}(L^{(+)}, L^{(-)}) \right). \end{aligned}$$

When there is no ambiguity, we omit the variable description $(L^{(+)}, L^{(-)})$ for Π_{\min} and Π_{sum} . Then the probabil-

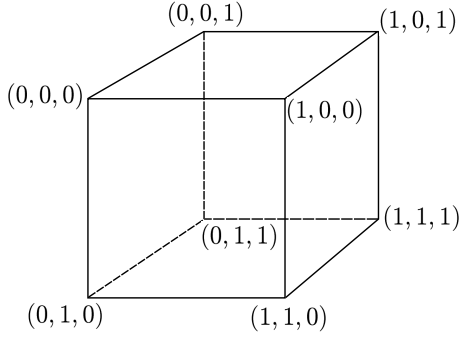


FIG. 4: A 3-dimensional cube. A random walk on the cube is equivalent to a Markov chain \mathcal{M} with $N = 3$.

ity distribution $\{\Pi_1(\mathbf{p}, \mathbf{q})\}$ is calculated to be

$$\begin{aligned} \Pi_1(\mathbf{p}, \mathbf{q}) &= \sum_{(\mathbf{p}', \mathbf{q}') \in \Sigma(L^{(+)}, L^{(-)})} \mathcal{G}[(\mathbf{p}, \mathbf{q}); (\mathbf{p}', \mathbf{q}')] \Pi_0(\mathbf{p}', \mathbf{q}') \\ &= \frac{1}{\Pi_{\text{sum}}} (\mathbb{E}[\xi_{T_W+1}(\mathbf{p}, \mathbf{q})] - \Pi_{\min}), \end{aligned}$$

where we use Proposition 4 and a fact that the matrix \mathcal{G} is bistochastic. Repeating this, the probability distribution after k steps is obtained by

$$\Pi_k(\mathbf{p}, \mathbf{q}) = \frac{1}{\Pi_{\text{sum}}} (\mathbb{E}[\xi_{T_W+k}(\mathbf{p}, \mathbf{q})] - \Pi_{\min}). \quad (15)$$

Thus we can define a Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$ on a l -dimensional hypercube with a transition matrix $\mathcal{G}[(\mathbf{p}, \mathbf{q}); (\mathbf{p}', \mathbf{q}')] and a probability distribution $\Pi_k(\mathbf{p}, \mathbf{q})$.$

Note that Eq. (15) leads to $\forall(\mathbf{p}, \mathbf{q}) \in \Sigma(L^{(+)}, L^{(-)})$

$$\left| \Pi_k(\mathbf{p}, \mathbf{q}) - \Pi_{\infty}(\mathbf{p}, \mathbf{q}) \right| = \frac{1}{\Pi_{\text{sum}}} \left| \mathbb{E}[\xi_{T_W+k}(\mathbf{p}, \mathbf{q})] - \mathbb{E}[\xi_{\infty}(\mathbf{p}, \mathbf{q})] \right|,$$

where $\Pi_{\infty}(\mathbf{p}, \mathbf{q})$ is a stationary distribution of the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$. This implies that if the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$ converges with an error $\epsilon/\Pi_{\text{sum}}$, $\mathbb{E}[\xi_{T_W}(\mathbf{p}, \mathbf{q})]$ converges with an error ϵ . Hence $T_{\text{conv}}(\epsilon) = T_{\text{mix}}(\epsilon/\Pi_{\text{sum}})$ where T_{mix} is the mixing time of the Markov chain defined in Appendix C.

The mixing time of the Markov chain on a hypercube depends on two factors. One is the dimension of the hypercube l and another is the probability that no change happens on a Markov chain, which is referred to as a *staying probability*. Obviously, a larger l and a smaller staying probability result in the longer mixing time. In the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$, the maximum of the dimension and the minimum of the staying probability are achieved for $l = N - 1$. If the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$ with $l = N - 1$ converges with an error $\epsilon/\Pi_{\text{sum}}$, the other Markov chains with $l \neq N - 1$ converges with an error less than $\epsilon/\Pi_{\text{sum}}$. Thus, hereafter, we consider only the Markov chain $\mathcal{M}(L^{(+)}, L^{(-)})$ with $l = N - 1$, which we denote by \mathcal{M} .

Finally, we simplify the notations of the Markov chain \mathcal{M} . Since it is equivalent to a Markov chain on a $(N - 1)$ -dimensional hypercube with a transition matrix \mathcal{G} , we label each vertex by a binary number $\vec{i} = (i_1, \dots, i_{N-1}) \in \{0, 1\}^{N-1}$, not by (\mathbf{p}, \mathbf{q}) . If we identify $(\mathbf{p}_0, \mathbf{q}_0)$ with $\vec{i}_0 = 0 \dots 0$, a new label for $f_k(\mathbf{p}_0, \mathbf{q}_0)$ is given by $0 \dots 010 \dots 0$ where only the k -th digit of the binary representation is 1. The vertices \vec{i} and \vec{j} are connected if and only if $H(\vec{i}, \vec{j}) = 1$ where $H(\vec{i}, \vec{j}) = \sum |i_k - j_k|$ (see Fig. 4). In this new labeling, the transition matrix $\mathcal{G}[(\mathbf{p}, \mathbf{q}); (\mathbf{p}', \mathbf{q}')] is simplified to$

$$\mathcal{P}(\vec{i}, \vec{j}) = \begin{cases} 1 - \frac{2}{N} & \text{if } \vec{i} = \vec{j}, \\ \frac{2}{N(N-1)} & \text{if } H(\vec{i}, \vec{j}) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

2. Lower bound of the mixing time of a Markov chain \mathcal{M}

The Markov chain on a $(N - 1)$ -dimensional hypercube with a staying probability $1/2$ has been well studied. For such a Markov chain, the transition matrix is given by $I/2 + \mathcal{P}'/2$ where \mathcal{P}' is a matrix of which elements are $\frac{1}{N-1}$ for $H(\vec{i}, \vec{j}) = 1$, and 0 otherwise. All eigenvalues of the transition matrix are known to be $\{1 - \frac{k}{N-1}\}_{k=0, \dots, N-1}$ [50]. On the other hand, it is observed from Eq. (16) that $\mathcal{P}(\vec{i}, \vec{j}) = (1 - \frac{2}{N})I + \frac{2}{N}\mathcal{P}'$. Hence, the eigenvalues of $\mathcal{P}(\vec{i}, \vec{j})$ are given by $\{1 - \frac{4k}{N(N-1)}\}_{k=0, \dots, N-1}$, which leads to $\max_{i=2,3, \dots} |\lambda_i| = 1 - \frac{4}{N(N-1)}$. Using Eq. (C1) in Appendix C, the mixing time $T_{\text{mix}}(\epsilon)$ is bounded from below by

$$\left(\frac{N(N-1)}{4} - 1 \right) \log \frac{1}{2\epsilon} \leq T_{\text{mix}}(\epsilon).$$

3. Upper bound of the mixing time of a Markov chain \mathcal{M}

In order to derive an upper bound of the mixing time, we slightly modify the transition matrix $\mathcal{P}(\vec{i}, \vec{j})$ by changing the staying probability from $1 - \frac{2}{N}$ to $1 - \frac{1}{N-1}$. A new transition matrix $\tilde{\mathcal{P}}(\vec{i}, \vec{j})$ is given by

$$\tilde{\mathcal{P}}(\vec{i}, \vec{j}) = \begin{cases} 1 - \frac{1}{N-1} & \text{if } \vec{i} = \vec{j}, \\ \frac{1}{(N-1)^2} & \text{if } H(\vec{i}, \vec{j}) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We denote by $\tilde{\mathcal{M}}$ a new Markov chain with a transition matrix $\tilde{\mathcal{P}}$. Since the staying probability of \mathcal{P} is smaller than $\tilde{\mathcal{P}}$, the mixing time of \mathcal{M} provides an upper bound of that of $\tilde{\mathcal{M}}$.

We investigate the mixing time of the Markov chain $\tilde{\mathcal{M}}$ by the *coupling method* (see Appendix C). For constructing a coupling, we interpret $\tilde{\mathcal{M}}$ as follows. At a step $t = k_0 + m(N - 1)$ ($k_0 = 1, \dots, N - 1$ and $m = 0, 1, \dots$),

choose $j \in \{i_1, \dots, i_{N-1}\}$ at random and, if $j = i_{k_0}$, flip i_{k_0} , otherwise, do nothing.

Based on this, we define a coupling (X_k, Y_k) . A state in X_k (Y_k) is a bit sequence $x_1 \dots x_{N-1}$ ($y_1 \dots y_{N-1}$) where x_j (y_j) $\in \{0, 1\}$. At a step $t = k_0 + m(N-1)$ ($k_0 = 1, \dots, N-1$ and $m = 0, 1, \dots$), we randomly choose $j \in \{i_1, \dots, i_{N-1}\}$. If $j \neq i_{k_0}$, do nothing. If $j = i_{k_0}$ and $x_j = y_j$, flip both of them. If $j = i_{k_0}$ and $x_j \neq y_j$, flip x_j and $y_{\alpha(j)}$ where $\alpha(j)$ is the position satisfying $x_{\alpha(j)} \neq y_{\alpha(j)}$ next to j . For instance, when $X = (0, 1, 0, 1, 1)$ and $Y = (1, 1, 0, 0, 0)$, $\alpha(1) = 4$ and $\alpha(4) = 5$. Each of X_k and Y_k is equivalent to the Markov chain $\tilde{\mathcal{M}}$ and they also satisfy $X_{k+1} = Y_{k+1}$ if $X_k = Y_k$. Hence, (X_k, Y_k) is a coupling of $\tilde{\mathcal{M}}$.

For investigating the coupling, we use the property of a special type of Markov chains, a *coupon collecting*. A coupon collecting of r -coupons is a Markov chain on a set of states $\{0, 1, \dots, r\}$ with the transition matrix given by

$$\begin{aligned} \text{Prob}[X_{t+1} = k+1 | X_t = k] &= \frac{r-k}{r}, \\ \text{Prob}[X_{t+1} = k | X_t = k] &= \frac{k}{r}. \end{aligned}$$

A coupon collecting of r coupons is interpreted as a trial of collecting a complete set of r different coupons by drawing one coupon at each step. When r_0 coupons are initially at hand, we denote the necessary number of steps to draw all coupons by $\tau_{\text{coupon}}^{(r, r_0)}$.

Proposition 6 Let (X_k, Y_k) be a coupling of $\tilde{\mathcal{M}}$ defined above and T_{xy} be its stopping time. Then T_{xy} is bounded from above by

$$T_{xy} \leq (N-1)\tau_{\text{coupon}}^{(N-1, N-2)}.$$

Proof 4 By definition of the coupling, for a fixed $k_0 \in \{1, \dots, N-1\}$, once i_{k_0} is chosen at the steps $k = k_0 + m(N-1)$ ($m = 0, 1, \dots$), we have $x_{i_{k_0}} = y_{i_{k_0}}$. The number of steps necessary for picking up k_0 from $\{1, \dots, N-1\}$ is equal to that of a coupon collecting of $(N-1)$ -coupons with initially $N-2$ coupons at hand, that is, $\tau_{\text{coupon}}^{(N-1, N-2)}$. Since there exists $N-1$ choices of k_0 , the stopping time T_{xy} is smaller than or equal to $(N-1)\tau_{\text{coupon}}^{(N-1, N-2)}$. ■

A coupon collecting of r -coupons starting with no coupon is a well-studied problem. It is known that if we draw coupons $r \log r$ times, we can complete all of r -coupons with a high probability, that is, $\tau_{\text{coupon}}^{(r, 0)}$ is typically given by $r \log r$. For $\tau_{\text{coupon}}^{(r, r_0)}$, we first show that

$$\text{Prob}[\tau_{\text{coupon}}^{(r, r_0)} > r(\log(r - r_0) + c)] \leq e^{-c}, \quad (16)$$

for any $c > 0$. This is shown in a standard way (see D for details). By appealing to Theorem 4 presented in

Appendix C, Proposition 6 and Eq. (16), we can bound $\Delta((N-1)^2 c)$ by

$$\begin{aligned} \Delta((N-1)^2 c) &\leq \max_{x, y} \text{Prob}[T_{xy} > (N-1)^2 c] \\ &\leq \text{Prob}[(N-1)\tau_{\text{coupon}}^{(N-1, N-2)} > (N-1)^2 c] \\ &= \text{Prob}[\tau_{\text{coupon}}^{(N-1, N-2)} > (N-1)c] \\ &< e^{-c}. \end{aligned}$$

Denote by $\tilde{T}_{\text{mix}}(\epsilon)$ the mixing time of the Markov chain $\tilde{\mathcal{M}}$ with an error ϵ . Since $\tilde{T}_{\text{mix}}(\epsilon)$ is defined by $\Delta(\tilde{T}_{\text{mix}}(\epsilon)) \leq \epsilon$, we obtain

$$\tilde{T}_{\text{mix}}(\epsilon) \leq (N-1)^2 \log \epsilon^{-1},$$

which also provides an upper bound of the mixing time of the Markov chain \mathcal{M} since $T_{\text{mix}}(\epsilon) < \tilde{T}_{\text{mix}}(\epsilon)$.

4. Upper and lower bounds of the convergence time $T_{\text{conv}}(\epsilon)$

We require an error $\epsilon/(\Pi_{\text{sum}} 2^{2N})$ in order to obtain an upper bound of the convergence time $T_{\text{conv}}(\epsilon)$ (see Eq. (13)) and, an error $\epsilon/\Pi_{\text{sum}}$ for a lower bound (see Eq. (14)). Recalling that the unitary operation \tilde{W} consists of $T_{\tilde{W}} = \lceil N/2 \rceil$ two-qubit gates, we obtain bounds for $T_{\text{conv}}(\epsilon)$ such that

$$\begin{aligned} \left(\frac{N(N-1)}{4} - 1\right) \log\left(\frac{\Pi_{\text{sum}}(L^{(+)}, L^{(-)})}{2\epsilon}\right) &\leq T_{\text{conv}}(\epsilon) - T_{\tilde{W}} \\ &\leq (N-1)^2 \log\left(\frac{2^{2N} \Pi_{\text{sum}}(L^{(+)}, L^{(-)})}{\epsilon}\right). \end{aligned}$$

Although $\Pi_{\text{sum}}(L^{(+)}, L^{(-)})$ depends on an initial state $|\phi_0\rangle$, it is shown that $\Pi_{\text{sum}}(L^{(+)}, L^{(-)}) \leq 2^N$ for any $(L^{(+)}, L^{(-)})$ and any $|\phi_0\rangle$. Therefore, we finally obtain

$$\begin{aligned} \frac{N^3}{4} \log 2 + \frac{N^2}{4} \log \epsilon^{-1} + O(N^2) &\leq T_{\text{conv}}(\epsilon) \\ &\leq 3N^3 \log 2 + N^2 \log \epsilon^{-1} + O(N^2), \end{aligned}$$

for any initial state $|\phi_0\rangle$. This concludes the proof of Theorem 2.

VI. SUMMARY AND CONCLUDING REMARKS

In this paper, we have introduced concepts of diagonal-unitary t -designs and toric t -designs that simulate up to the t -th order of statistical moments of diagonal-unitary matrices and phase-random states, respectively. We have presented how to implement diagonal-unitary 2-designs in the computational basis for N -qubit systems by using two types of the phase-random circuits, the CP and the CZ phase-random circuit. We have shown that the CP phase-random circuit exactly achieves a diagonal-unitary

2-design after applying $O(N^2)$ up to two-qubit diagonal gates. On the other hand, the CZ phase-random circuit approximately achieves a diagonal-unitary 2-design after applying $O(N^2(N + \log 1/\epsilon))$ two-qubit diagonal gates, showing that random variables in the genuine two-qubit diagonal gate provide stronger ability for randomizing phases. Since they are composed of only diagonal gates and, for the CP phase-random circuit, all the gates can be applied simultaneously, it is expected that the phase-random circuits are practically easy to implement.

We have also provided applications of the phase-random circuits. One is a generation of complex-projective 2-design by combining a phase-random circuit with a simple classical procedure. Compared to the previously known results, our method is simple for implementation in the sense that all gates are diagonal. Another is a quantum simulator of canonical states of classical Hamiltonians where canonical states with any temperature can be implemented by preparing a proper initial state, without realizing the Hamiltonian itself.

In analogy with the random circuits, which are shown to approximately achieve unitary t -designs for any t by applying $\text{poly}(N, t)$ two-qubit gates [31], it is natural to guess that the phase-random circuits with appropriate gate sets would also achieve diagonal-unitary t -designs in $\text{poly}(N, t)$ iterations. However, it is not the case if we use the gate sets composed of only two-qubit diagonal gates since there is a lack of the number of parameters due to the commutability of gates. Thus, for constructing diagonal-unitary t -designs, the gate set should include multi-qubit gates if only diagonal gates are used. It would be interesting to specify the diagonal gate set of the phase-random circuit achieving diagonal-unitary t -designs, or to construct a quantum circuit composed of non-diagonal two-qubit gates which achieves diagonal-unitary t -designs.

Acknowledgement

The authors thank F. G. S. L. Brandao for insightful comments on unitary t -designs and P. S. Turner for help-

ful discussions. This work was supported by Project for Developing Innovation Systems of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. Y. N. acknowledges support from JSPS by KAKENHI (Grant No. 222812) and M. M. acknowledges support from JSPS by KAKENHI (Grant No. 23540463).

Appendix A: Calculation of $G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$

$G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$ is defined by

$$G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') = \mathbb{E}[\text{Tr} \sigma_{\mathbf{p}} W_{ij} \sigma_{\mathbf{p}'} W_{ij}^\dagger \text{Tr} \sigma_{\mathbf{q}} W_{ij} \sigma_{\mathbf{q}'} W_{ij}^\dagger],$$

where W_{ij} is a two-qubit diagonal gate on the i -th and j -th qubits and randomly chosen from the gate set $\mathcal{W}^{CZ} = \{\text{diag}(1, e^{i\alpha}, e^{i\beta}, -e^{i(\alpha+\beta)})\}_{\alpha, \beta}$ or $\mathcal{W}^{CP} = \{\text{diag}(1, e^{i\alpha}, e^{i\beta}, e^{i\gamma})\}_{\alpha, \beta, \gamma}$. Since \mathcal{W}^{CP} is more general than \mathcal{W}^{CZ} , we start with the calculation of $G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$ for \mathcal{W}^{CP} .

In order to calculate $G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}')$, we define \mathcal{D}_{ab} , \mathcal{E}_{ab} and Δ_a by

$$\begin{aligned} \mathcal{D}_{ab} &= \delta_{a0}\delta_{bz} + \delta_{az}\delta_{b0}, \\ \mathcal{E}_{ab} &= \delta_{ax}\delta_{by} - \delta_{ay}\delta_{bx}, \end{aligned}$$

and

$$\Delta_a = \delta_{a0} + \delta_{az} - \delta_{ax} - \delta_{ay}.$$

We also use a notation that $\delta_{n \in S} = 1$ if $n \in S$ and $\delta_{n \in S} = 0$ if $n \notin S$.

For $W_{ij} = \text{diag}(1, e^{i\alpha}, e^{i\beta}, e^{i\gamma})$, it is straightforward to calculate $\text{Tr} \sigma_{\mathbf{p}} W_{ij} \sigma_{\mathbf{p}'} W_{ij}^\dagger$ and we obtain

$$\begin{aligned}
\frac{1}{2N} \text{Tr} \sigma_{\mathbf{P}} W_{ij} \sigma_{\mathbf{P}'} W_{ij}^\dagger = & \delta_{\mathbf{P}, \mathbf{P}'} \left\{ \delta_{p_i, p_j \in \{0, z\}} + \frac{1}{2} \delta_{p_i \in \{x, y\}} (\cos \beta + \cos(\alpha - \gamma)) + \frac{1}{2} \delta_{p_j \in \{x, y\}} (\cos \alpha + \cos(\beta - \gamma)) \right. \\
& \left. - \frac{1}{2} \delta_{p_i, p_j \in \{x, y\}} [\cos \alpha + \cos \beta - \cos \gamma - \cos(\alpha - \beta) + \cos(\alpha - \gamma) + \cos(\beta - \gamma)] \right\} \\
& + \frac{1}{2} \delta_{p_i, p'_i} \left\{ \delta_{p_i \in \{x, y\}} (\cos \beta - \cos(\alpha - \gamma)) \mathcal{D}_{p_j, p'_j} \right. \\
& \left. - [\delta_{p_i \in \{0, z\}} (\sin \alpha - \sin(\beta - \gamma)) + \delta_{p_i \in \{x, y\}} (\sin \gamma + \sin(\alpha - \beta))] \mathcal{E}_{p_j, p'_j} \right\} \\
& + \frac{1}{2} \delta_{p_j, p'_j} \left\{ \delta_{p_j \in \{x, y\}} (\cos \alpha - \cos(\beta - \gamma)) \mathcal{D}_{p_i, p'_i} \right. \\
& \left. - [\delta_{p_j \in \{0, z\}} (\sin \beta - \sin(\alpha - \gamma)) + \delta_{p_j \in \{x, y\}} (\sin \gamma + \sin(\beta - \alpha))] \mathcal{E}_{p_i, p'_i} \right\} \\
& - \frac{1}{2} (\sin \beta + \sin(\alpha - \gamma)) \mathcal{E}_{p_i, p'_i} \mathcal{D}_{p_j, p'_j} - \frac{1}{2} (\sin \alpha + \sin(\beta - \gamma)) \mathcal{D}_{p_i, p'_i} \mathcal{E}_{p_j, p'_j} \\
& + \frac{1}{2} (\cos(\alpha - \beta) - \cos \gamma) \mathcal{E}_{p_i, p'_i} \mathcal{E}_{p_j, p'_j}. \tag{A1}
\end{aligned}$$

In the case of \mathcal{W}^{CP} , by taking the average over $\alpha, \beta, \gamma = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$, G_{ij} is calculated to

$$\begin{aligned}
\frac{1}{2^{2N}} G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') = & \delta_{\mathbf{p}, \mathbf{p}'} \delta_{\mathbf{q}, \mathbf{q}'} \left(\delta_{p_i, p_j, q_i, q_j \in \{0, z\}} + \frac{1}{4} \delta_{p_i, q_i \in \{0, z\}} \delta_{p_j, q_j \in \{x, y\}} + \frac{1}{4} \delta_{p_i, q_i \in \{x, y\}} \delta_{p_j, q_j \in \{0, z\}} + \frac{1}{4} \delta_{p_i, p_j, q_i, q_j \in \{x, y\}} \right) \\
& + \frac{1}{4} \delta_{p_i, q_i \in \{x, y\}} \delta_{p_i, p'_i} \delta_{q_i, q'_i} \left(\mathcal{D}_{p_j, p'_j} \mathcal{D}_{q_j, q'_j} + \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j} \right) + \frac{1}{4} \delta_{p_j, q_j \in \{x, y\}} \delta_{p_j, p'_j} \delta_{q_j, q'_j} \left(\mathcal{D}_{p_i, p'_i} \mathcal{D}_{q_i, q'_i} + \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \right) \\
& + \frac{1}{4} \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \left(\delta_{p_j, q_j \in \{0, z\}} \delta_{p_j, p'_j} \delta_{q_j, q'_j} + \mathcal{D}_{p_j, p'_j} \mathcal{D}_{q_j, q'_j} \right) + \frac{1}{4} \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j} \left(\delta_{p_i, q_i \in \{0, z\}} \delta_{p_i, p'_i} \delta_{q_i, q'_i} + \mathcal{D}_{p_i, p'_i} \mathcal{D}_{q_i, q'_i} \right) \\
& + \frac{1}{4} \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j}.
\end{aligned}$$

By investigating each case, we obtain Eq. (4).
On the other hand, in the case of \mathcal{W}^{CZ} , γ is set to be

$\alpha + \beta + \pi$ and G_{ij} is obtained as

$$\begin{aligned}
G_{ij}(\mathbf{p}, \mathbf{q}; \mathbf{p}', \mathbf{q}') = & 2^{2N} \left[\delta_{\mathbf{p}, \mathbf{p}'} \delta_{\mathbf{q}, \mathbf{q}'} \left(\delta_{p_i, p_j, q_i, q_j \in \{0, z\}} + \frac{1}{4} \delta_{p_i, p_j, q_i, q_j \in \{x, y\}} \right) \right. \\
& + \frac{1}{2} \delta_{p_i, q_i \in \{x, y\}} \delta_{p_i, p'_i} \delta_{q_i, q'_i} \left(\mathcal{D}_{p_j, p'_j} \mathcal{D}_{q_j, q'_j} + \frac{1}{2} \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j} \right) + \frac{1}{2} \delta_{p_j, q_j \in \{x, y\}} \delta_{p_j, p'_j} \delta_{q_j, q'_j} \left(\mathcal{D}_{p_i, p'_i} \mathcal{D}_{q_i, q'_i} + \frac{1}{2} \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \right) \\
& \left. + \frac{1}{2} \mathcal{D}_{p_i, p'_i} \mathcal{D}_{q_i, q'_i} \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j} + \frac{1}{2} \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \mathcal{D}_{p_j, p'_j} \mathcal{D}_{q_j, q'_j} + \frac{1}{4} \mathcal{E}_{p_i, p'_i} \mathcal{E}_{q_i, q'_i} \mathcal{E}_{p_j, p'_j} \mathcal{E}_{q_j, q'_j} \right],
\end{aligned}$$

leading to Eq. (6).

Appendix B: Calculation of the expectation of the state after the phase-random circuits

Here, we show that

$$2^{-N} \left[\sum_{S_{\text{even}}(\mathbf{p}, \mathbf{q})} - \sum_{S_{\text{odd}}(\mathbf{p}, \mathbf{q})} \right] \xi_0(\mathbf{p}', \mathbf{q}') = \mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]. \tag{B1}$$

For an initial state of the phase-random circuit represented by $|\phi_0\rangle = \sum r_n e^{i\omega_n} |\bar{n}\rangle$, $\xi_0(\mathbf{p}, \mathbf{q})$ is given by $\xi_0(\mathbf{p}, \mathbf{q}) = 2^{-N} \langle \phi_0 | \sigma_{\mathbf{p}} | \phi_0 \rangle \langle \phi_0 | \sigma_{\mathbf{q}} | \phi_0 \rangle$. A straightforward calculation shows that

$$\begin{aligned} \xi_0(\mathbf{p}, \mathbf{q}) &= 2^{-N} \sum_{n,m,l,k} r_n r_m r_l r_k e^{i(\omega_n - \omega_m + \omega_l - \omega_k)} \\ &\times \prod_{j \in \Lambda^{(+)}} \delta_{n_j m_j} \delta_{l_j k_j} [\delta_{p_j 0} + \delta_{p_j z} (1 - 2n_j)(1 - 2l_j)] \\ &\times \prod_{j \in \Lambda^{(-)}} \delta_{n_j m_j} \delta_{l_j k_j} [\delta_{p_j 0} (1 - 2l_j) + \delta_{p_j z} (1 - 2n_j)] \\ &\times \prod_{j \in \Gamma^{(+)}} \bar{\delta}_{n_j m_j} \bar{\delta}_{l_j k_j} [\delta_{p_j x} - \delta_{p_j y} (1 - 2n_j)(1 - 2l_j)] \\ &\times \prod_{j \in \Gamma^{(-)}} i(\bar{\delta}_{n_j m_j} \bar{\delta}_{l_j k_j} [\delta_{p_j x} (1 - 2l_j) + \delta_{p_j y} (1 - 2n_j)]), \end{aligned}$$

where $\bar{\delta}_{nm} = 1 - \delta_{nm}$. By substituting $\xi_0(\mathbf{p}, \mathbf{q})$ into the left-hand side of Eq. (B1), we obtain

$$\begin{aligned} \mathbb{E}[\xi_\infty(\mathbf{p}, \mathbf{q})] &= 2^{-2N} \sum_{n,m,l,k} r_n r_m r_l r_k e^{i(\omega_n - \omega_m + \omega_l - \omega_k)} \\ &\times \prod_{j \in \Lambda^{(+)}} 2\delta_{n_j m_j} \delta_{l_j k_j} \delta_{n_j l_j} \prod_{j \in \Lambda^{(-)}} 2\delta_{n_j m_j} \delta_{l_j k_j} \delta_{n_j l_j} (1 - 2n_j) \\ &\times \prod_{j \in \Gamma^{(+)}} 2\delta_{n_j k_j} \delta_{m_j l_j} \bar{\delta}_{n_j m_j} \\ &\times \prod_{j \in \Gamma^{(-)}} -2i\delta_{n_j k_j} \delta_{m_j l_j} (\delta_{p_j x} - \delta_{p_j y}) \bar{\delta}_{n_j m_j} (1 - 2n_j), \end{aligned}$$

where we have used a relationship that

$$\begin{aligned} &\left[\sum_{S_{\text{even}}(\mathbf{p}, \mathbf{q})} - \sum_{S_{\text{odd}}(\mathbf{p}, \mathbf{q})} \right] \\ &\prod_{j \in \Gamma^{(-)}} i\bar{\delta}_{n_j m_j} \bar{\delta}_{l_j k_j} [\delta_{p_j x} (1 - 2l_j) + \delta_{p_j y} (1 - 2n_j)] \\ &= \prod_{j \in \Gamma^{(-)}} -2i\bar{\delta}_{n_j m_j} \bar{\delta}_{l_j k_j} (\delta_{p_j x} - \delta_{p_j y})(n_j - l_j) \\ &= \prod_{j \in \Gamma^{(-)}} -2i\delta_{n_j k_j} \delta_{m_j l_j} (\delta_{p_j x} - \delta_{p_j y}) \bar{\delta}_{n_j m_j} (1 - 2n_j). \end{aligned}$$

Finally, using $\lambda^{(+)} + \lambda^{(-)} + \gamma^{(+)} + \gamma^{(-)} = N$, we obtain the result such that

$$\begin{aligned} &2^{-N} \left[\sum_{S_{\text{even}}(\mathbf{p}, \mathbf{q})} - \sum_{S_{\text{odd}}(\mathbf{p}, \mathbf{q})} \right] \xi_0(\mathbf{p}', \mathbf{q}') = \\ &2^{-N} \sum_{n,m} r_n^2 r_m^2 \prod_{j \in \Lambda^{(+)}} \delta_{n_j m_j} \prod_{j \in \Lambda^{(-)}} \delta_{n_j m_j} (1 - 2n_j) \\ &\times \prod_{j \in \Gamma^{(+)}} \bar{\delta}_{n_j m_j} \prod_{j \in \Gamma^{(-)}} -i\bar{\delta}_{n_j m_j} (\delta_{p_j x} - \delta_{p_j y})(1 - 2n_j), \end{aligned}$$

which is equal to $\mathbb{E}_{\mathcal{U}_{\text{diag}}^{(2)}} [\xi_\varphi(\mathbf{p}, \mathbf{q})]$ given by Eq. (3).

Appendix C: Introduction of Markov chain

A Markov chain is a sequence of random variables indexed by a discrete *step* $t \in \mathbb{N}$ that take values in a set of *states* $S = \{s\}$. We define a probability distribution $\{\Pi_t(s)\}_{s \in S}$ at a step t over the state space S . The Markov property is that a probability distribution Π_{t+1} depends only on Π_t . This evolution of the probability distribution is governed by a stochastic *transition matrix* \mathcal{P} such that $\Pi_{t+1} = \mathcal{P}\Pi_t$. The element of a transition matrix is denoted by $\mathcal{P}(s, s')$, which represents a probability that a transition from s to s' occurs. Using an initial distribution Π_0 , the probability distribution at step t is given by $\Pi_t = \mathcal{P}^t \Pi_0$. A Markov chain is said to be irreducible (aperiodic) when a transition matrix is irreducible (aperiodic). For an irreducible and aperiodic Markov chain, the Perron-Frobenius theorem guarantees that there exists a unique stationary distribution $\Pi_\infty = \lim_{t \rightarrow \infty} \Pi_t$ independent of the initial probability distribution.

We define the *mixing time*. The mixing time is the number of steps required for the actual distribution to be close to the stationary distribution, where the distance after t -steps is defined by

$$\Delta(t) := \max_{s \in S} |\Pi_t(s) - \Pi_\infty(s)|.$$

The mixing time $T_{\text{mix}}(\epsilon)$ is defined by for any $\epsilon > 0$,

$$T_{\text{mix}}(\epsilon) := \min\{t | \Delta(t') \leq \epsilon \text{ for all } t' \geq t\}.$$

In order to study an upper bound and a lower bound of the mixing time, we introduce a *relaxation time* of a Markov chain. Denote the eigenvalues of a transition matrix \mathcal{P} by λ_i ($i = 1, 2, \dots$) in decreasing order. When a transition matrix is irreducible and aperiodic, it is known that $1 = \lambda_1 > \lambda_2$. A *relaxation time* T_{rel} is defined by

$$T_{\text{rel}} = (1 - \max_{i=2,3,\dots} |\lambda_i|)^{-1},$$

which gives bounds of the mixing time $T_{\text{mix}}(\epsilon)$ such that

$$(T_{\text{rel}} - 1) \log\left(\frac{1}{2\epsilon}\right) \leq T_{\text{mix}}(\epsilon) \leq \log\left(\frac{1}{\epsilon \Pi_{\min}}\right) T_{\text{rel}}, \quad (\text{C1})$$

where $\Pi_{\min} := \min_s \Pi_\infty(s)$ is the minimum stationary distribution [50].

Although the relaxation time provides both of the upper and the lower bounds of the mixing time, it does not give tight bounds. Hence, we introduce a *coupling* method for investigating the upper bound of the mixing time. A pair of two random walks (X_t, Y_t) , where t denotes the number of steps, is said to be a coupling of a Markov chain when the following two conditions are satisfied. First, each of X_t and Y_t is a faithful copy of the Markov chain. Second, (X_t, Y_t) should satisfy the condition that $X_t = Y_t$ implies $X_{t+1} = Y_{t+1}$. For a coupling (X_t, Y_t) , we define a *stopping time* T_{xy} by

$$T_{xy} := \min\{t | X_t = Y_t, \text{ when } X_0 = x, Y_0 = y\}.$$

By definition, $X_t = Y_t$ for all $t > T_{xy}$. The stopping time is related to the mixing time through the following theorem [50].

Theorem 4 Let (X_t, Y_t) be a coupling of a Markov chain and T_{xy} be the stopping time. Then,

$$\Delta(t) \leq \max_{x,y} \text{Prob}[T_{xy} > t].$$

Since the mixing time is obtained from $\Delta(t)$, we can derive an upper bound of the mixing time from the stopping time.

In the following, we use a relaxation time to obtain a lower bound of the mixing time and investigate an upper bound of the mixing time by using the coupling method.

Appendix D: Coupon collecting starting with non-zero coupons

We consider a coupon collecting and show that for $c > 0$,

$$\text{Prob}[\tau_{\text{coupon}}^{(r,r_0)} > r(\log(r - r_0) + c)] \leq e^{-c}. \quad (\text{D1})$$

Suppose that any of the j -th coupons ($j = 1, \dots, r - r_0$) are initially not at hand. We denote by A_j an event where the j -th coupon has not been collected within $r(\log(r - r_0) + c)$ steps. An upper bound of the left hand side of Eq. (D1) is obtained by

$$\begin{aligned} & \text{Prob}[\tau_{\text{coupon}}^{(r,r_0)} > r(\log(r - r_0) + c)] \\ &= \text{Prob}[\cup_{j=1}^{r-r_0} A_j] \\ &\leq \sum_{j=1}^{r-r_0} \text{Prob}[A_j] \\ &= \sum_{j=1}^{r-r_0} (1 - \frac{1}{r})^{r(\log(r-r_0)+c)} \\ &= (r - r_0)(1 - \frac{1}{r})^{r(\log(r-r_0)+c)} \\ &\leq (r - r_0) \exp[-\log(r - r_0) - c] \\ &= e^{-c}. \end{aligned}$$

-
- [1] M. L. Metha, *Random Matrices*, Academic Press (1990).
 - [2] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor and A. Winter, *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp 56-74 (2005).
 - [3] B. M. Terhal, David P. DiVincenzo and D. W. Leung, *Phys. Rev. Lett.* 86, 5807-5810 (2001).
 - [4] D. P. DiVincenzo and D. W. Leung and B. M. Terhal, *IEEE Trans. Inf Theory* Vol. 48. No. 3, 580-599 (2002).
 - [5] P. O. Boykin, V. Roychowdhury, *Phys. Rev. A* 67, 042317 (2003).
 - [6] P. Hayden, D. Leung, P. W. Shor and A. Winter *Comm. Math. Phys.* 250(2): 371-391 (2004).
 - [7] P. Sen, *IEEE Conf. on Computational Complexity*, pp 274-287 (2006).
 - [8] C. Dankert, R. Cleve, J. Emerson and E. Livine, *Phys. Rev. A*, 80, 012304 (2009).
 - [9] A. Abeyesinghe, I. Devetak, P. Hayden and A. Winter, *Proc. R. Soc. A* 465, pp 2537-2563 (2009).
 - [10] P. Hayden and J. Preskill, *JHEP* 0709:120 (2007).
 - [11] F. Dupuis, M. Berta, J. Wullschleger and R. Renner, *arXiv:1012.6044* (2010).
 - [12] E. Lubkin, *J. Math. Phys.* 19 1028 (1978).
 - [13] D. N. Page, *Phys. Rev. Lett.* 71 1291, (1993).
 - [14] S. K. Foong and S. Kanno, *Phys. Rev. Lett.* 72, 1148 (1994).
 - [15] J. S. Ruiz, *Phys. Rev. E* 52, 5653-5655 (1995).
 - [16] K. Życzkowski and H.J. Sommers, *J. Phys. A: Math. Gen.* 34 (2001) 7111-7125.
 - [17] P. Hayden, D. W. Leung and A. Winter, *Comm. Math. Phys.* Vol. 265, No. 1, pp. 95-117 (2006).
 - [18] D. Gross, S. T. Flammia and J. Eisert, *Phys. Rev. Lett.* 102, 190501 (2009).
 - [19] S. Goldstein, J. L. Lebowitz, R. Tumulka and N. Zanghi *Phys. Rev. Lett.* 96, 050403 (2006).
 - [20] S. Popescu, A. J. Short and A. Winter, *Nature Physics*, 2:754-758 (2006).
 - [21] P. Reimann, *Phys. Rev. Lett.* 101, 190403 (2008).
 - [22] N. Linden, S. Popescu, A. J. Short and A. Winter, *Phys. Rev. E* 79, 061103 (2009).
 - [23] A. Riera, C. Gogolin and J. Eisert, *Phys. Rev. Lett.* 108, 080402 (2012).
 - [24] Vinayak and M. Žnidarić, *J. Phys. A: Math. Gen.* 45 (2012) 125204.
 - [25] L. Masanes, A. J. Roncaglia and A. Acín, *arXiv:1108.0374* (2011).
 - [26] F. G. S. L. Brandao, P. Ćwikliński, M. Horodecki, P. Horodecki, J. Korbicz and M. Mozzyrmas, *Phys. Rev. E* 86, 031101 (2012).
 - [27] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd and D. G. Cory. *Science*, 302, pp 2098-2100 (2003).
 - [28] A. W. Harrow and R. A. Low, *Proceedings of RANDOM 2009*, LNCS, 5687:548-561 (2009).
 - [29] A. W. Harrow and R. A. Low, *Comm. Math. Phys.* Vol. 291, No. 1, pp. 257-302 (2009); I. T. Diniz and D. Jonathan, *Comm. Math. Phys.* Vol. 304, No 1, pp 281-293 (2011).
 - [30] F. G. S. L. Brandao and M. Horodecki, *arXiv:1010.3654* (2010).
 - [31] F. G.S.L. Brandao, A. W. Harrow, M. Horodecki, *arXiv:1208.0692*, (2012).
 - [32] Y. Nakata, P. S. Turner and M. Muraio, *Phys. Rev. A* 86, 012301 (2012).
 - [33] P. Delsartem, J. M. Goethals and J. J. Seidel, *Geom. Dedicata* 6 (1977), 363-388.
 - [34] P. D. Seymour and T. Zaslarsky, *Advances in Mathematics* 52, 213-240 (1984).

- [35] H. Barnum, arXiv:quant-ph/0205155, (2002).
- [36] J. M. Renes, R. Blume-Kohout, A. J. Scott and C. M. Caves, J. Math. Phys., 45, 6, (2004).
- [37] A. Ambainis and J. Emerson, IEEE Conference on Computational Complexity page 129-140, 2007 (2007).
- [38] A. Roy and A. J. Scott, J. Math. Phys. 48, 072110 (2007).
- [39] A. Y. Kitaev, Russian Math. Surveys 52, 1191 (1997).
- [40] R. A. Low, Proceeding of the Royal Society A, 465, 2111 3289-3308.
- [41] A. Hayashi, T. Hachimoto and M. Horibe Phys. Rev. A, **72** 032325 (2006).
- [42] D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, Springer-Verlag, New York (1986).
- [43] S. Aaronson, Proceedings of IEEE Conference on Computational Complexity, pages 229-242, 2009 (2009).
- [44] R. Oliveira, O. C. O. Dahlsten and M. B. Plenio, Phys. Rev. Lett., 98, 130502 (2007); O. C. O. Dahlsten, R. Oliveira and M. B. Plenio, J. Phys. A: Math. and Theor., 40, pp 8081-8108, (2007).
- [45] A. Kitaev, arXiv:quant-ph/9511026.
- [46] N. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, England, (2000).
- [47] K. Temme, T. J. Osborne, K. Vollbrecht, D. Poulin and F. Verstraete, Nature 471, 87 (2011).
- [48] D. Poulin and P. Wocjan, Phys. Rev. Lett. 103, 220502 (2009).
- [49] R. G. Horn and C. R. Johnson, *MATRIX ANALYSIS*, Cambridge university Press (1985).
- [50] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, Providence (2009).